

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

**Defective images within this document are accurate representations of
the original documents submitted by the applicant.**

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORLED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 54
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-09-815-626-6

Query Match 29.2%; Score 83; DB 10; Length 54;
Best Local Similarity 35.3%; Pred. No. 0.0018;
Matches 18; Conservative 7; Mismatches 22; Indels 4; Gaps 1;

QY 1 NPMACDCRARPPLMAMFQ----RARVSSSDVTCATPPEROGRLRLREADF 47
DB 1 NPFNCDCELRWLLRWLRRTNPRLEDEDLRCASPESLRLGQPLLELPSEDF 51

RESULT 3
Sequence 6, Application US/09822687
Patent No. US20020076753A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria Alexandra
TITLE OF INVENTION: 31939, A NOVEL HUMAN LEUCINE-RICH REPEAT
FILE REFERENCE: 10448-038001
CURRENT APPLICATION NUMBER: US/09/822,687
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,919
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 54
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-09-822-687-6

Query Match 29.2%; Score 83; DB 10; Length 54;
Best Local Similarity 35.3%; Pred. No. 0.0018;
Matches 18; Conservative 7; Mismatches 22; Indels 4; Gaps 1;

QY 1 NPMACDCRARPPLMAMFQ----RARVSSSDVTCATPPEROGRLRLREADF 47
1 NPFNCDCELRWLLRWLRRTNPRLEDEDLRCASPESLRLGQPLLELPSEDF 51

RESULT 4
Sequence 6, Application US/09789404
Patent No. US2002002554A1
GENERAL INFORMATION:
APPLICANT: Khodadoust, Mehrian
TITLE OF INVENTION: NOVEL LEUCINE RICH REPEAT-CONTAINING MOLECULES AND USES THEREFOR
FILE REFERENCE: 10448/008001
CURRENT APPLICATION NUMBER: US/09/789,404
CURRENT FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 09/456,592
PRIOR FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 54
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-09-789-404-6

Query Match 28.2%; Score 80; DB 10; Length 54;
Best Local Similarity 35.3%; Pred. No. 0.0042;
Matches 18; Conservative 6; Mismatches 23; Indels 4; Gaps 1;

QY 1 NPMACDCRARPPLMAMFQ----RARVSSSDVTCATPPEROGRLRLREADF 47
DB 1 NPFNCDCELRWLLRWLRRTNPRLEDEDLRCASPESLRLGQPLLELPSEDF 51

RESULT 5
Sequence 12, Application US/09815626
Patent No. US20020076752A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria Alexandra
TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT
FILE REFERENCE: 10448-031001
CURRENT APPLICATION NUMBER: US/09/815,626
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,863
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 55
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-09-815-626-12

Query Match 25.9%; Score 73.5; DB 10; Length 55;
Best Local Similarity 30.8%; Pred. No. 0.027;
Matches 16; Conservative 9; Mismatches 22; Indels 5; Gaps 1;

QY 1 NPMACDCRARPPLMAMFQARVSS-----SDVTCATPPEROGRLRLREADF 47
DB 1 NPFNCDCELRWLLRWLRLEAQNNEALODPVSLRCASPESLRLGQPLLELPSEF 52

RESULT 6
Sequence 977, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 977
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-300-977

Query Match 20.1%; Score 57; DB 10; Length 59;
Best Local Similarity 43.5%; Pred. No. 3;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 PMACDCRARPPLMAMFQARVSSS 24
DB 37 PFCGCRARSVSWFCPAMIRFS 59

RESULT 7
US-09-764-877-1630
Sequence 1630, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1630
LENGTH: 86
TYPE: PRF
ORGANISM: Homo sapiens
US-09-764-877-1630

Query Match 18.0%; Score 51; DB 10; Length 86;
at Local Similarity 41.7%; Pred. No. 25;
Matches 10; Conservative 0; Mismatches 8; Indels 6; Gaps 1;

Oy 3 WACDCRRAR-----PLAMFORAR 20
Db 60 WAOCRATACSVHPPSWGWRQMR 83

RESULT 8
US-09-739-907-98
Sequence 98, Application US/09739907
Patent No. US20010012889A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P202281
CURRENT APPLICATION NUMBER: US/09/739,907
CURRENT FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,692
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,704
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,658
PRIOR FILING DATE: 1998-01-07
NUMBER OF SEQ ID NOS: 196
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 98
LENGTH: 62
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (62)
OTHER INFORMATION: Xaa equals stop translation
US-09-739-907-98

Query Match 17.3%; Score 49; DB 10; Length 62;
Best Local Similarity 28.6%; Pred. No. 30;
Matches 14; Conservative 6; Mismatches 13; Indels 16; Gaps 2;

Oy 9 ARPLAMFORARV-----SSSDVTCARP-----PERGRDLA 41
Db 4 ARPPMAYLALITALLLGTEHYLVANDVSCDHSNTVPSSMDLGA 52

RESULT 9
US-09-764-877-1144
Sequence 1144, Application US/09764877
Patent No. US20020147140A1

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1144
LENGTH: 91
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (25)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (27)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (31)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (35)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (36)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (54)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (57)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1144

Query Match 16.9%; Score 48; DB 10; Length 91;
Best Local Similarity 37.5%; Pred. No. 61;
Matches 15; Conservative 6; Mismatches 11; Indels 8; Gaps 2;

Oy 7 GRAPPLAMFORARVSSSDVTCARP-----PERGRDL 39
Db 48 CLARPRXGQXQASRYVSMN-GCASPQHDHSAKHDPDKCNL 86

RESULT 10
US-10-001-883-123
Sequence 123, Application US/10001883
Publication No. US20030022188A1
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herve
APPLICANT: Pluta, Jason
APPLICANT: Ghosh, Malayika
APPLICANT: Sun, YONGMING
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and
FILE REFERENCE: DEX-0271
CURRENT APPLICATION NUMBER: US/10/001,883
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/252,059
PRIOR FILING DATE: 2000-11-20
NUMBER OF SEQ ID NOS: 137
SOFTWARE: Patentln version 3.1
SEQ ID NO 123
LENGTH: 42
TYPE: PRF
ORGANISM: Homo sapien
US-10-001-883-123

Query Match 16.7%; Score 47.5; DB 9; Length 42;
 Best Local Similarity 45.8%; Pred. No. 30;
 Matches 11; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

OY 11 PLWAFORA-RVSSSDVTCATPPE 33
 DB 7 PLTSFTTRALRYKASSMCLAPPE 30

RESULT 11
 US-09-864-761-42876

; Sequence 42876, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aecm1ca-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 42876
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AF161326.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
 ; OTHER INFORMATION: EST HUMAN HIT: A1632488.1, EVALUATE 3.00e-21
 ; OTHER INFORMATION: SWISSPROT HIT: P05998, EVALUATE 3.10e-01
 ; US-09-864-761-42876

Query Match 16.5%; Score 47; DB 10; Length 50;
 Best Local Similarity 33.3%; Pred. No. 42;
 Matches 15; Conservative 5; Mismatches 17; Indels 8; Gaps 2;

OY 3 WACDRAPLMAM---FORARVSSSDVTCATPPEGRDLRLRE 44
 DB 4 WSLTWAPLTMAMVPPFGGPRATSPMPCGWPGRK---RALLE 43

RESULT 12
 US-09-950-933A-46

; Sequence 46, Application US/09950933A
 ; Patent No. US2002016141A1
 ; GENERAL INFORMATION:

; APPLICANT: Simmons, Carl R.
 ; APPLICANT: Navarro, Pedro
 ; TITLE OF INVENTION: Antimicrobial Peptides and Methods of
 ; FILE REFERENCE: Use
 ; CURRENT APPLICATION NUMBER: US/09/950,933A
 ; CURRENT FILING DATE: 2001-09-11
 ; PRIOR APPLICATION NUMBER: 60/232,569
 ; PRIOR FILING DATE: 2000-09-13
 ; NUMBER OF SEQ ID NOS: 99
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 77
 ; TYPE: PRT
 ; ORGANISM: Trifolium aestivum
 ; US-09-950-933A-46

Query Match 16.2%; Score 46; DB 9; Length 77;
 Best Local Similarity 29.3%; Pred. No. 89;
 Matches 12; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

OY 2 PWACD--CARPLMFORARVSSSDVTCAT---PPEROG 36
 DB 16 FWECSKSSRCSGTGYKACLTGCKCATCLVPPGTYG 56

RESULT 13
 US-09-864-761-38710

; Sequence 38710, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
 ; FILE REFERENCE: Aecm1ca-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38710
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006548.19
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: P50343, EVALU 9.20e-01
US-09-864-761-38710

Query Match          15.8%; Score 45; DB 10; Length 53;
Best Local Similarity 38.5%; Pred. No. 79;
Matches 10; Conservative 3; Mismatches 5; Indels 8; Gaps 2;

OY 2 PMA---CDCRARIWAMQARVSS 24
Db 32 PWSVSLSDSTCKPPWAM-----ISSS 52

; Sequence 1003, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR APPLICATION NUMBER: 2001-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
```

```

; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1003
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1003

Query Match          15.8%; Score 45; DB 9; Length 61;
Best Local Similarity 35.7%; Pred. No. 92;
Matches 10; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

OY 7 GRARPLWAMF---QRAVSSSDVTCAT 30
Db 2 CFPRLMSMASPTPSSATLRTSSHQCSF 29

RESULT 15
US-09-864-761-47370
; Sequence 47370, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47370
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL137225.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.45
US-09-864-761-47370

```

```

; Every Match 15.38; Score 43.5; DB 10; Length 52;
; st Local Similarity 33.38; Pred. No. 1.2e+02;
; Matches 13; Conservative 4; Mismatches 17; Indels 5; Gaps 1;
QY 12 LMWQBARAVSSDVTCATPPROGDRALREADPQAC 50
| | : | | | | | | | | | | | | | | | |
Db 1 LMCSSKRAKPS-----LLMFGPERSLRLVMPKARRSCC 34

```

Search completed: April 4, 2003, 08:31:56
Job time : 85.5556 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 42.963 Seconds

(without alignments)
155,076 Million cell updates/sec

Title: US-09-972-546-2_COPY_261_310

Perfect score: 284
Sequence: 1 NFWACDCRRARPLWAFORAR.....PPERGRDLRALREADFOAC 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Number of hits satisfying chosen parameters: 180751

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : A_Geneseq.101002:*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	62.3	51	22	AAB87100
2	100	58.8	51	22	AAB87115
3	100	58.8	51	22	AAO21485
4	91	32.0	49	22	AAE03627
5	91	32.0	49	22	AAB82379
6	83	29.2	54	22	AAE13009
7	83	29.2	54	22	AAE65808
8	83	29.2	54	22	AAE03603
9	83	29.2	54	22	AAE23805
10	80	28.2	54	22	AAB82355

11	73.5	25.9	55	22	AAE65814	33395 LRR domain c
12	62	21.8	75	22	AAU64229	Protonibacterium
13	62	21.8	99	22	AAU65387	Protonibacterium
14	57	20.1	43	12	AAE14369	Amorphous domain o
15	57	20.1	59	21	AAE56399	Human prostate can
16	57	20.1	86	22	ABG26420	Novel human diagno
17	56	19.7	99	22	AAU47800	Protonibacterium
18	53.5	18.8	52	22	AAU49313	Protonibacterium
19	53.5	18.8	90	22	AAU39937	Protonibacterium
20	53	18.7	47	22	ABG21804	Novel human diagno
21	53	18.7	63	22	AAU61143	Protonibacterium
22	52	18.3	80	21	AAE38009	Human secreted pro
23	52	18.3	60	20	AAV11476	Human 5' EST secre
24	52	18.3	91	23	ABE08809	Human protein kin
25	51	18.0	69	23	ABP08373	Human ORFX protein
26	51	18.0	86	22	ABP09568	Human musculoskele
27	51	18.0	86	22	ABE03683	Human polyepitide
28	51	18.0	95	22	AAO07825	Human polyepitide
29	50.5	17.8	53	22	AAU56161	Protonibacterium
30	50.5	17.8	53	22	AAU66070	Protonibacterium
31	50.5	17.8	84	21	AAE00181	Human secreted pro
32	50.5	17.8	93	22	AAE28193	Protonibacterium
33	50	17.6	58	22	AAU49947	Protonibacterium
34	49.5	17.4	53	22	AAU49947	Protonibacterium
35	49	17.3	62	20	AAV38427	Human secreted pro
36	49	17.3	67	22	AAU53141	Protonibacterium
37	49	17.3	94	22	AAU57157	Protonibacterium
38	48.5	17.1	60	22	AAU50568	Protonibacterium
39	48.5	17.1	65	22	AAU45258	Protonibacterium
40	48.5	17.1	83	22	AAU63604	Protonibacterium
41	48	16.9	41	22	AAE03946	Human gene 49 enco
42	48	16.9	50	22	AAU48483	Protonibacterium
43	48	16.9	51	22	AAE84924	Human immune/haema
44	48	16.9	61	22	AAU64105	Protonibacterium
45	48	16.9	69	21	AAE40541	Human ORFX ORF305

ALIGNMENTS

RESULT 1	
1	AAE87100
ID	AAE87100 standard; Protein; 51 AA.
AC	AAE87100;
XX	
XX	
DT	04-MAY-2001 (first entry)
XX	
DE	Human TANGO 393 Leu-rich-repeat domain, SEQ ID NO: 88.
XX	
KW	Secreted protein; transmembrane protein; TANGO; human; drug screening;
KW	activity modulator; expression modulator; cancer; immunological disorder;
KW	cytostatic; immunomodulatory; gene therapy.
OS	Homo sapiens.
XX	
PN	WO200109162-A2.
XX	
PD	08-FEB-2001.
XX	
PF	31-JUL-2000; 2000WO-US20935.
XX	
PR	30-JUL-1999; 99US-0365164.
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS;
PI	Pan Y;
XX	
DR	WPI; 2001-138647/14.
XX	
PT	Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
	for the prevention, diagnosis and treatment of, e.g. cancers and immune

PT disorders -
XX
PS Disclosure: Page 253; 332pp; English.
XX
XX The invention relates to novel secreted/transmembrane proteins, and
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
CC TANGO 333, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
CC murine TANGO 393 is also included within the scope of the invention. The
CC invention also encompasses fragments and variants of the proteins of the
CC invention, and nucleic acids encoding them. The invention additionally
CC relates to host cells comprising a nucleic acid of the invention; methods
CC for the production of a protein of the invention; an antibody specific
CC for a protein of the invention; methods for detecting a protein or
CC nucleic acid of the invention; and methods of identifying agents which
CC bind to or modulate the activity of a protein of the invention. The novel
CC secreted proteins, nucleic acids encoding them, and antibodies against
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression or activity of the secreted
CC proteins. The secreted proteins of the invention may also be used to
CC identify modulators of expression or activity, which may be useful in
CC the treatment of disorders associated with the proteins of the
CC invention e.g., cancers and immunological disorders. The present
CC sequence represents a fragment of a human TANGO protein of the
CC invention.
SQ Sequence 51 AA:
Query Match 62.3%; Score 177; DB 22; Length 51;
Best Local Similarity 60.0%; Pred. No. 8.7e-16;
Matches 30; Conservative 4; Mismatches 16; Indels 0; Gaps 0;
OY 1 NPMACDCRARPPLMAMFORARVSSSDVTGATPPERQGRDLALREADPQAC 50
DB 1 NPMWCDCRARPPLMAMLOKFRGSSSEVPCNLPQRLAGDLRLAANDLQGC 50
RESULT 2
AAB87115
ID AAB87115 standard; Protein; 51 AA.
XX
AC AAB87115;
XX
DT 04-MAY-2001 (first entry)
XX
DE Mouse TANGO 393 Leu-rich-repeat domain, SPQ ID NO:108.
XX
KW Secreted protein; transmembrane protein; TANGO 393; mouse; murine;
drug screening; activity modulator; expression modulator; cancer;
immunological disorder; cytostatic; immunomodulatory; gene therapy.
OS Mus musculus.
XX
XX WO200109162-A2.
XX
XX 08-FEB-2001.
XX
XX 31-JUL-2000; 2000WO-US20935.
XX
XX 30-JUL-1999; 99US-0365164.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Fraser CC, Sharp JD, KIRST SJ, Barnes TM, WRIGHTON N, MYERS PS;
PI Pan Y;
XX
XX WPI; 2001-138647/14.
XX
XX Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
PT for the prevention, diagnosis and treatment of, e.g. cancers and immune
PT disorders -
XX
PS Disclosure: Page 266; 332pp; English.

XX
CC The invention relates to novel secreted/transmembrane proteins, and
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
CC TANGO 333, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
CC murine TANGO 393 is also included within the scope of the invention. The
CC invention also encompasses fragments and variants of the proteins of the
CC invention, and nucleic acids encoding them. The invention additionally
CC relates to host cells comprising a nucleic acid of the invention; methods
CC for the production of a protein of the invention; an antibody specific
CC for a protein of the invention; methods for detecting a protein or
CC nucleic acid of the invention; and methods of identifying agents which
CC bind to or modulate the activity of a protein of the invention. The novel
CC secreted proteins, nucleic acids encoding them, and antibodies against
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression or activity of the secreted
CC proteins. The secreted proteins of the invention may also be used to
CC identify modulators of expression or activity, which may be useful in
CC the treatment of disorders associated with the proteins of the
CC invention e.g., cancers and immunological disorders. The present
CC sequence represents a fragment of the mouse TANGO 393 protein of the
CC invention.
SQ Sequence 51 AA:
Query Match 58.8%; Score 167; DB 22; Length 51;
Best Local Similarity 56.0%; Pred. No. 1.8e-14;
Matches 28; Conservative 5; Mismatches 17; Indels 0; Gaps 0;
OY 1 NPMACDCRARPPLMAMFORARVSSSDVTGATPPERQGRDLALREADPQAC 50
DB 1 NPMWCDCRARPPLMAMLOKFRGSSSEVPCNLPQRLADRLRLAASDLQGC 50
RESULT 3
AA021485
ID AA021485 standard; Protein; 50 AA.
XX
XX AA021485;
XX
DT 15-AUG-2002 (first entry)
XX
XX Consensus NGR LRCT domain protein sequence.
XX
XX Cerebroprotective; neuroprotective; cytostatic; Nogo receptor homolog;
KW NGR2; NGR3; axonal growth; central nervous system; CNS; cerebral injury;
KW spinal cord injury; stroke; demyelinating disease; multiple sclerosis;
KW monophasic demyelination; encephalomyelitis; Marchiafava-Bignami disease;
KW multifocal leukoencephalopathy; panencephalitis; Spongy degeneration;
KW Alexander's disease; Canavan's disease; metachromatic leukodystrophy;
KW Krabbe's disease; immune; bait protein; genetic mapping; gene therapy;
KW transgenic animal; unregulated cellular growth; cancer; tumour.
XX
XX Unidentified.
XX
XX WO200229059-A2.
XX
XX 11-APR-2002.
XX
XX 06-OCT-2001; 2001WO-US31488.
XX
XX 06-OCT-2000; 2000US-238361P.
XX
XX (UYVA) UNIV YALE.
XX
XX (BIJU) BIOGEN INC.
XX
XX Strittmatter SM, Cate RL, Sah DWY;
PI
XX
XX WPI; 2002-416677/44.
XX
XX Novel Nogo receptor homolog polypeptide, NGR2 or NGR3, useful for
PT treating central nervous system disorder, cerebral injury, spinal cord
PT injury, stroke, and demyelinating diseases -
XX
PT

XX Claim 13: Page 107: 277pp: English.

CC The invention relates to a Nogo receptor homologue polypeptide, NGR2 or
CC NGR3, comprising a 50 amino acid LRRCT sequence, a 284 amino acid NTLRRCT
CC sequence, or a 420, 461 or 392 amino acid sequence, all given in the
CC specification. The NGR3 protein or its binding antibody is useful for
CC decreasing inhibition of axonal growth of a central nervous system (CNS)
CC neuron, by contacting the neuron NGR3 or its antibody, and for treating
CC CNS disease, disorder or injury. NGR3 or a vector comprising NGR3 is
CC useful for treating cerebral injury, spinal cord injury, stroke,
CC demyelinating diseases, e.g. multiple sclerosis, monophasic
CC demyelination, encephalomyelitis, multifocal leukoencephalopathy,
CC paraneoplasia, Marchiafava-Bignami disease, Spongy degeneration,
CC Alexander's disease, Canavan's disease, metachromatic leukodystrophy and
CC Krabbe's disease. NGR3 is useful for inducing an immune response in a
CC mammal against NGR3, as a bait protein in a two-hybrid or three-hybrid
CC assay, and as a research tool for identification, characterisation and
CC purification of interacting, regulatory proteins. The nucleotide
CC sequences of the invention are useful for screening for RFLP associated
CC with certain disorders, for genetic mapping, and for gene therapy. The
CC vector containing NGR3 is useful for producing non-human transgenic
CC animals. The NGR3 binding antibody is useful for isolating and purifying
CC NGR3, for localisation and/or quantitation of NGR3, and for diagnostic
CC and therapeutic purposes. The sequences of the invention, vectors and
CC antibodies are useful for treating or preventing unregulated cellular
CC growth such as cancer and tumour growth. This sequence represents the
CC consensus NGR LRRCT domain of the invention.

XX Sequence 50 AA:

Query Match 44.4%; Score 126; DB 23; Length 50;
Best Local Similarity 46.0%; Pred. No. 4e-09;
Matches 23; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

OY 1 NPMACDCRRPLMAMFORARVSSDYTCATPPEROGRDLRALREADFOAC 50
Db 1 NNMKXCRRARLMMWXXXKSSXXVXXCXXGDXLXXLXXDXXC 50

RESULT 4

AE03627 standard: peptide: 49 AA.

AC AE03627;

DT 07-AUG-2001 (first entry)

Human AZAD protein C-terminal leucine-rich repeat sequence.

KW Human; leucine-rich repeat; LRR; AZAD; neurodegenerative disorder;
KW CNS disorder; central nervous system disorder; prostate disorder;
KW prostatic; benign prostatic hyperplasia; adenocarcinoma;
KW prostate cancer; genitourinary system carcinoma; testicular tumour;
KW Alzheimer's disease; dementia; Parkinson's disease; multiple sclerosis;
KW amyotrophic lateral sclerosis; epilepsy; psychiatric disorder; mania;
KW depression; schizophrenia; anxiety; phobic disorder; learning disorder;
KW memory disorder; amnesia; migraine; protein-protein interaction;
KW cellular activity; neoplastic transformation; drug screening;
KW forensic identification; gene therapy.

OS Homo sapiens.

XX MO200142286-A2.

XX 14-JUN-2001.

XX 07-DEC-2000; 2000MO-US31140.

XX 08-DEC-1999; 99US-0456592.

XX (MILL-) MILLENNIUM PHARM INC.

PI Khodadoust MM;

DR WPI: 2001-381633/40.

PT New leucine-rich repeat-containing protein and nucleic acid molecules
PT for diagnosing, treating neural disorders, such as neurodegenerative
PT disorders, such as Alzheimer's disease, dementia, epilepsy and prostate
PT cancer
XX Example 1: Fig 4; 133pp: English.

CC The present sequence is the C-terminal leucine-rich repeat of
CC human AZAD protein which is a multiple leucine-rich repeat-containing
CC secreted protein. AZAD polypeptide is useful for
CC identifying a compound which modulates its activity and binds to it.
CC AZAD proteins and nucleic acid molecules are useful for treating and
CC diagnosing AZAD-mediated or related disorders, which includes a neural
CC disorder (e.g. neurodegenerative disorders, including CNS disorders)
CC and a prostate disorder (e.g. prostatic, benign prostatic hyperplasia
CC and cancer (e.g. adenocarcinoma, prostate cancer, genitourinary system
CC carcinomas and testicular tumour). The neurodegenerative disorders
CC include Alzheimer's disease, dementias related to Alzheimer's disease,
CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
CC epilepsy, psychiatric disorders e.g. depression, schizophrenia
CC disorders, mania, anxiety or phobic disorders, learning or memory
CC disorders e.g. amnesia or age-related memory loss, and neurological
CC disorders e.g. migraine. AZAD molecules are useful as markers of
CC disorders or disease states, as markers for precursors of disease state,
CC for predisposition of disease state or as markers of drug activity or
CC pharmacogenomic profile of a subject. The AZAD polypeptide and
CC polynucleotide are capable of modulating protein-protein interaction,
CC e.g. by interacting with an extracellular component, thereby modulating
CC cellular activities, including attachment, adhesion, migration,
CC patterning, growth and/or differentiation of a cell. AZAD proteins
CC regulate embryonic development and differentiation, tissue maintenance
CC and function, pathological conditions, e.g. neuronal degeneration,
CC neoplastic transformation and tumour progression. AZAD proteins are
CC useful as immunogens to raise anti-AZAD antibodies which are useful to
CC detect and isolate AZAD proteins and modulate its activity. AZAD
CC proteins are useful to screen for naturally occurring AZAD substrates
CC and to screen for drugs or compounds which modulate AZAD activity.
CC AZAD nucleic acid fragments are useful as primers or hybridisation
CC probes for the detection of AZAD-encoding nucleic acids. cDNA encoding
CC the AZAD protein are useful in gene therapy. AZAD sequences are also
CC useful to map their respective genes on a chromosome, for tissue typing
CC and in forensic identification of a biological sample.

XX Sequence 49 AA:

Query Match 32.0%; Score 91; DB 22; Length 49;
Best Local Similarity 40.4%; Pred. No. 0.00015;
Matches 21; Conservative 8; Mismatches 17; Indels 6; Gaps 3;

OY 1 NPMACDCRRPLMAMFO--RARVSSDYTCATPPEROGRDLRALREADFOAC 50
Db 1 NPFHDCOLLPLHRLTGLNLRVGA---TCATPPMARQGRVKA-AAAVFEDC 48

RESULT 5

AAB82379 standard: Protein: 49 AA.

XX AAB82379;

XX 23-JUL-2001 (first entry)

XX Protein sequence SEQ ID NO. 30.

XX JAFFA; human; fibroblast growth factor.

XX Homo sapiens.

XX WO200138357-A2.

XX 31-MAY-2001.
 XX
 PD 22-NOV-2000; 2000WO-US32181.
 XX
 PF 22-NOV-1999; 9905-0444165.
 XX
 PR (MILL-) MILENNIUM PHARM INC.
 XX
 PA
 PI Khodadoust MM;
 XX
 DR WPI; 2001-355881/37.
 XX
 XX Isolated JAFFA nucleic acid molecules which encode novel fibroblast
 PT growth factor family members (JAFFA) are useful for developing novel
 PT diagnostic and therapeutic agents for JAFFA-associated disorders such
 PT as cancers -
 XX
 XX Disclosure; Page 136; 137pp; English.
 XS

The present invention relates to human JAFPA polynucleotides (see AAR90326) and polypeptides (see AAR82251). JAFPA is a novel member of the fibroblast growth factor family. JAFPA proteins may be useful for developing novel diagnostic and therapeutic agents for JAFPA-associated disorders such as cancers and ABOH blood group disorders, and for controlling cellular proliferative and/or differentiative disorders. JAFPA nucleic acids and proteins may be used to treat and/or diagnose a variety of immune disorders such as autoimmune disease and multiple sclerosis. JAFPA nucleic acids may also be used to express JAFPA protein, to detect JAFPA mRNA or a genetic alteration in a JAFPA gene, and to modulate JAFPA activity. Note: The present sequence is given in the Sequence Listing from the present invention, but does not correspond with any SEQ ID NO. mentioned within the specification.

50 Sequence 49 AA;

Query Match	32.0%;	Score 91;	DB 22;	Length 49;
Best Local	40.4%;	Pred. No. 0.00015;		
Matches 21;	Conservative	8;	Mismatches 17;	Indels 6;
				Gaps 3

```
Qy      1 NPWACDCRRAPLMAMTQ--RARVSSSDVTCA TPPEPQGRLRALREADIQAC 50
        || : ||| | | : ||||| :| : :| | :| :|
Db      1 NPFHCQCQLPLRHKWL TGLNLRVGA----TCATPPNARGRVKA-AAAYFEDC 48
```

RESULT 6
AAE13009

AAE13009;

28-JAN-2002 (first entry)

Consensus C-terminal leucine-rich repeat (LRRCT) domain.

KM Human; leucine-rich repeat: LRR: 31939 protein therapy;
KM cell proliferation; differentiation disorder; cancer; neuronal disorder;
KM neurological disorder; demyelinating disease; multiple sclerosis;
KM degenerative disease; Alzheimer's disease; Huntington's disease;
KM spinocerebellar degeneration; nervous system; bone disorder;
KM osteoporosis; immune disorder; rheumatoid arthritis; diabetes mellitus;
KM cardiovascular disorder; liver disorder; viral disease; pain;
KM metabolic disorder; chromosomal mapping; tissue typing; forensic biology;
KM cytostatic; neurotropic; neuroprotective; anticonvulsant; osteopathic;
KM antirheumatic; antiarthritis; vinclole; analgesic.

Homo sapiens.

W0200175105-A2.

11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US10380.

PR 31-MAR-2000; 2000US-193919P.

PA (MILL-) MILLENNIUM PHARM INC.
PA
PA

PI Glucksmann MA;

DR WPI; 2001-662974/76.

PT New polypeptide useful for treating cancer, multiple sclerosis,
PT Alzheimer's disease, osteoporosis, arthritis, and metabolic and liver
PT disorders comprises the 31939 polypeptide belonging to the leucine-rich
PT repeat family -

PS Disclosure; Fig 3; 117pp; English.

The present sequence is a consensus human C-terminal leucine-rich repeat (LRR) domain. The LRR family member (referred as 3199) protein is useful for treating cell proliferative or differentiation disorder, e.g., cancer disorders or neuronal disorders, neurological disorders, demyelinating diseases such as multiple sclerosis, degenerative diseases such as Alzheimer's disease and Huntington's disease, spinocerebellar degenerations, disorders of the central or peripheral nervous system, bone disorders such as osteoporosis, immune disorders such as rheumatoid arthritis, diabetes mellitus, cardiovascular disorders, liver disorders, viral diseases, pain or metabolic disorders. They are also useful as markers of disorders or disease states, or for precursors or markers for the predisposition of disease states. They are useful in screening and detection assays such as chromosomal mapping, tissue typing, and forensic biology.

SQ Sequence 54 AA;

Query Match	29.28;	Score 83;	DB 22;	length 54;
Best Local Similarity	35.38;	Pred. No. 0.0018;		
Matches 18; Conservative	7;	Mismatches 22;	Indels 4;	Gaps 1;

```
Qy 1 NPWACDCRAPPLIMAFQ---KARVSSSDVTCATPPEROGDRLRALREADF 47
   ||: ||| | | : | | | | : ||
Db 1 NPFNCDCELRWLLRWLRETNPRALEDQEDLRCAESPESLRGQPLELLPSDF 51
```

RESULT 7
AAG65808

XX
AC AAG65808;

DT 30-JAN-2002 (first entry)

DE 33395 LRR domain corresponding consensus amino acid fragment

KM LRR: leucine rich repeat, 33335; cytostatic; anti-HIV; antidiabetic;
 KM antiarthritic; neuroprotective; dermatological; immunosuppressive;
 KM antiinflammatory; antiasthmatic; antiulcer; antianaemic; hepatotropic

05 Homo sapiens.

PN WO200172827-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WO-US09470.

PR 24-MAR-2000; 2000US-191863P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Glucksmann MA;

DR WPI; 2001-626254/72.

XX


```
PT New polypeptide for preventing or treating disorders associated with
PT cellular adhesion, proliferation or differentiation, comprises
PT polypeptide 33395, a member of the leucine rich repeat protein family
PS -
XX Disclosure: Fig 3A-B; 133pp; English.
CC The invention provides an isolated nucleic acid encoding a polypeptide
CC of the leucine rich repeat (LRR) family, designated 33395. The 33395
CC polypeptide can be expressed by standard recombinant methodology. The
CC 33395 polynucleotides and polypeptide can be used to prevent or treat
CC disorders associated with 33395 expression, for example those involving
CC aberrant cellular adhesion, proliferation or differentiation. Specific
CC examples include melanoblastoma, juvenile AIDS, diabetes mellitus,
CC rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis,
CC Crohn's disease, ulcerative colitis, asthma, anemia, and chronic active
CC hepatitis. Sequences AAG65806-814 represent consensus amino acid
CC sequences derived from a hidden Markov model corresponding to the LRR
CC domain of the the human 33395 polypeptide.
CC Sequence 54 AA:
YY
DB 1 NPMACDCRARPPLMAWFO---RAVSSSDVMCAFPENOGRLRALREADF 47
   ||| ||| | | | : | | | | | | | | | | | | | | |
DB 1 NPFNCDCLRWLRWLRETNPRLDQEDLRKASPESLRGCPLELLLSDF 51
RESULT 8
AAE03603
ID AAE03603 standard; Protein; 54 AA.
XX
XX AAE03603:
XX
XX 07-AUG-2001 (first entry)
XX
XX Leucine-rich repeat consensus #3.
DE
KW Human; leucine-rich repeat; LRR; AZAD; neurodegenerative disorder;
KW CNS disorder; central nervous system disorder; prostate disorder;
KW prostatitis; benign prostatic hyperplasia; adenocarcinoma;
KW prostate cancer; genitourinary system carcinoma; testicular tumor;
KW Alzheimer's disease; dementia; Parkinson's disease; multiple sclerosis;
KW amyotrophic lateral sclerosis; epilepsy; psychiatric disorder; mania;
KW depression; schizophrenia; anxiety; phobic disorder; learning disorder;
KW memory disorder; amnesia; migraine; protein-protein interaction;
KW cellular activity; neoplastic transformation; drug screening;
KW forensic identification; gene therapy.
XX
XX Unidentified.
OS
XX WO200142286-A2.
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-US33140.
XX
XX 08-DEC-1999; 99US-0456592.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Khodadoust MM;
PI
XX WPI: 2001-381633/40.
XX
XX New leucine-rich repeat-containing protein and nucleic acid molecules
XX for diagnosing, treating neural disorders, such as neurodegenerative
XX disorders, such as Alzheimer's disease, dementia, epilepsy and prostate
XX cancer -
```

XX Example 1; Fig 4; 133bp; English.

CC The present sequence is a consensus sequence that

CC shares homology with the C-terminal leucine-rich repeat sequence of

CC human AZAD protein which is a multiple leucine-rich repeat-containing

CC secreted protein. AZAD polypeptide is useful for

CC identifying a compound which modulates its activity and binds to it.

CC AZAD proteins and nucleic acid molecules are useful for treating and

CC diagnosing AZAD-mediated or related disorders, which includes a neural

CC disorder (e.g. neurodegenerative disorders, including CNS disorders)

CC and a prostate disorder, e.g. prostatitis, benign prostatic hyperplasia

CC and cancer (e.g. adenocarcinoma, prostate cancer, genitourinary system

CC carcinomas and testicular tumour). The neurodegenerative disorders

CC include Alzheimer's disease, dementias related to Alzheimer's disease,

CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,

CC epilepsy, psychiatric disorders e.g. depression, schizophrenia

CC disorders, mania, anxiety or phobic disorders, learning or memory

CC disorders e.g. amnesia or age-related memory loss, and neurological

CC disorders e.g. migraine. AZAD molecules are useful as markers of

CC disorders or disease states, as markers for precursors of disease state,

CC for predisposition of disease state or as markers of drug activity or

CC pharmacogenomic profile of a subject. The AZAD polypeptide and

CC polynucleotide are capable of modulating protein-protein interaction,

CC e.g. by interacting with an extracellular component, thereby modulating

CC cellular activities, including attachment, adhesion, migration,

CC patterning, growth and/or differentiation of a cell. AZAD proteins

CC regulate embryonic development and differentiation, tissue maintenance

CC and function, pathological conditions, e.g. neuronal degeneration,

CC neoplastic transformation and tumour progression. AZAD proteins are

CC useful as immunogens to raise anti-AZAD antibodies which are useful to

CC detect and isolate AZAD proteins and modulate its activity. AZAD

CC proteins are useful to screen for naturally occurring AZAD substrates

CC and to screen for drugs or compounds which modulate AZAD activity.

CC AZAD nucleic acid fragments are useful as primers or hybridisation

CC probes for the detection of AZAD-encoding nucleic acids. cDNA encoding

CC the AZAD protein are useful in gene therapy. AZAD sequences are also

CC useful to map their respective genes on a chromosome, for tissue typing

CC and in forensic identification of a biological sample.

XX

SQ Sequence 54 AA:

OY Query Match 29.2%; Score 83; DB 22; Length 54;
Best Local Similarity 35.3%; Pred. No. 0.0018;
Matches 18; Conservative 7; Mismatch 22; Indels 4; Gaps 1;

Dy 1 NPMACDCARPLMAWFO---RARSSSDVYCAMPPEOGNDRLRLRADRF 47
||:||||| | : | : ||::||: ||: | | : ||
ID AAE23805 standard; Protein: 54 AA.
AAE23805

RESULT 9
AAE23805

XX AAE23805:
AC
DE Fibrinogen-like (NOVS) protein conserved LRR C-terminal (URCT) domain.
XX
XX 10-SEP-2002 (first entry)
DT
XX

KM Membrane bound protein; secreted NOV protein; spermatogenesis; neoplasia;
KM male infertility; angiogenesis; vascular pathology; orthopaedic disorder;
KM inflammatory disease; congenital muscular dystrophy; muscular disorder;
KM rheumatoid arthritis; fixed deformity; dysprolithrombinemia; cancer;
KM arthropodipositis; hypoprothrombinemia; hypokalaemic period paralysis;
KM Smith-Lemli-Oppitz syndrome; carcinoma tumour; centrocyclic lymphoma;
KM hyperparathyroidism; Leigh syndrome; cervical carcinoma; leukaemia;
KM macular dystrophy; vitelliform type; McRide disease; Meckel syndrome;
KM multiple endocrine neoplasia I; multiple myeloma; hyperparathyroidism;
KM parathyroid adenomatosis I; prolactinoma; digenic retinitis pigmentosa;
KM somatotrophinoma; neovascular inflammatory vitreoretinopathy; arthritis;
KM carcinoid syndrome; atopy; tendinitis; gene therapy; vaccine; URCT;
KM leucine rich repeat C-terminal domain; fibrinogen-like protein.

XX OS unidentified.
XX PN WO200230979-A2.
XX PD 18-APR-2002.
XX PF 10-OCT-2001; 2001WO-US31498.
XX PR 12-OCT-2000; 2000US-0689486.
XX PR 13-OCT-2000; 2000US-0687276.
XX PR 09-OCT-2001; 2001US-0973424.
XX PA (CURA-) CURAGEN CORP.
XX PI Prayaga SK, Taupier RJ, Bandaru R;
XX DR WPI; 2002-452341/48.
XX
XX Novel membrane bound and secreted NOV polypeptides, for treating,
XX diagnosing and preventing male infertility, neurological, cardiac and
XX vascular pathologies, and inflammatory diseases e.g. rheumatoid
XX arthritis -
XX
XX Disclosure: Page 38; 180pp; English.
XX
XX The present invention relates to novel membrane bound and secreted NOV
XX proteins and polynucleotides encoding such proteins. Sequences of the
XX invention are useful for treating or preventing NOV-associated disorders
XX in humans and for manufacturing a medicament for treating a syndrome
XX associated with human disease. They are useful for determining the
XX presence of or predisposition to lung cancer. NOV1 compounds are useful
XX for development, differentiation and activation of thymic immune cells,
XX pathologies related to spermatogenesis and male infertility, diagnosis
XX of several human neoplasias and diseases or pathologies of cells in
XX blood circulation such as red blood cells and platelets. NOV1 nucleic
XX acids are useful for detecting specific cell types and as specific
XX marker for cancers in tissues. NOV2 and NOV4 compounds are useful to
XX direct the development of nervous system and angiogenesis and for
XX treating neurological, cardiac and vascular pathologies. NOV3 and NOV5
XX compounds are useful for treating various orthopaedic disorders and/or
XX injuries, inflammatory diseases of connective tissue e.g. rheumatoid
XX arthritis, congenital muscular dystrophies, various muscular disorders,
XX fixed deformities (arthrogryposis) and abnormal white matter. They are
XX useful for treating atrophy, dysprothrombinaemia, hypoprothrombinaemia,
XX type I and type II Smith-Lemli-Opitz syndrome, carcinoid tumour of lung,
XX centrocystic lymphoma, cervical carcinoma, hyperparathyroidism, leish
XX syndrome, hypokalaemic periodic paralysis, acute promyelocytic leukaemia,
XX NDM4/RNA type, macular dystrophy, vitelliform type, Mcardle disease,
XX type 2 Meckel syndrome, multiple endocrine neoplasia I, multiple myeloma,
XX parathyroid adenomatosis I, prolactinoma, hyperparathyroidism, carcinoid
XX syndrome, digenic retinitis pigmentosa, somatotrophinoma, neovascular
XX inflammatory vitreoretinopathy, arthritis and tendonitis. Sequences of
XX the invention are also used in gene therapy and as vaccines. The present
XX sequence is fibromodulin-like (NOV5) protein conserved leucine rich
XX repeat C-terminal (LRRCT) domain.
XX
XX Sequence 54 AA:
XX
XX Query Match 29.2%; Score 83; DB 23; Length 54;
XX Best Local Similarity 35.3%; Pred. No. 0.0018;
XX Matches 18; Conservative 7; Mismatches 22; Indels 4; Gaps 1;
XX
XX 1 NPMACDCRRARPLAMFQ----RARVSSDPVTCATPPEGRGRRLRLREADF 47
XX ||: ||| | | | : | : ||: | : | | | : ||
XX Db 1 NPFNCDCELRLMLRLMRLTNPRLRLEDEGDLRCASPESLRGQPLLELPSDF 51

RESULT 10
AAB82355
ID AAB82355 standard; Protein: 54 AA.
XX
XX AAB82355;
XX
XX

XX 23-JUL-2001 (first entry)
XX DE Protein sequence SEQ ID NO.6.
XX
XX JAFFA; human; fibroblast growth factor.
XX KW
XX OS Homo sapiens.
XX PN WO200138357-A2.
XX PD 31-MAY-2001.
XX PF 22-NOV-2000; 2000WO-US32181.
XX PR 22-NOV-1999; 99US-0444165.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Khodadoust MM;
XX DR WPI; 2001-355881/37.
XX
XX Isolated JAFFA nucleic acid molecules which encode novel fibroblast
XX growth factor family members (JAFFA) are useful for developing novel
XX diagnostic and therapeutic agents for JAFFA-associated disorders such
XX as cancers -
XX
XX Disclosure: Page 132; 137pp; English.
XX
XX The present invention relates to human JAFFA polynucleotides (see
XX AAF0326) and polypeptides (see AAB82351). JAFFA is a novel member
XX of the fibroblast growth factor family. JAFFA proteins may be
XX useful for developing novel diagnostic and therapeutic agents for
XX JAFFA-associated disorders such as cancers and ABO(H) blood group
XX disorders, and for controlling cellular proliferative and/or
XX differentiative disorders. JAFFA nucleic acids and proteins may be
XX used to treat and/or diagnose a variety of immune disorders such as
XX autoimmune disease and multiple sclerosis. JAFFA nucleic acids may
XX also be used to express JAFFA protein, to detect JAFFA mRNA or a
XX genetic alteration in a JAFFA gene, and to modulate JAFFA activity.
XX Note: The present sequence is given in the Sequence listing from
XX the present invention, but does not correspond with the SEQ ID NO.
XX described within the specification. It is described in the
XX Sequencing listing as a consensus sequence.
XX
XX Sequence 54 AA:
XX
XX Query Match 28.2%; Score 80; DB 22; Length 54;
XX Best Local Similarity 35.3%; Pred. No. 0.0044;
XX Matches 18; Conservative 6; Mismatches 23; Indels 4; Gaps 1;
XX
XX 1 NPMACDCRRARPLAMFQ----RARVSSDPVTCATPPEGRGRRLRLREADF 47
XX ||: ||| | | | : | : ||: | : | | | : ||
XX Db 1 NPFNCDCELRLMLRLMRLTNPRLRLEDEGDLRCASPESLRGQPLLELPSDF 51

RESULT 11
AAG65814
ID AAG65814 standard; Protein: 55 AA.
XX
XX AAG65814;
XX
XX 30-JAN-2002 (first entry)
XX DE 33395 LRR domain corresponding consensus amino acid fragment.
XX
XX LRR; leucine rich repeat; 33395; cytosstatic; anti-HIV; antidiabetic;
XX antiairchritic; neuroprotective; dermatological; immunosuppressive;
XX antiinflammatory; antiasthmatic; antulcer; antianaemic; hepatotropic.
XX
XX Homo sapiens.
XX
XX

```

XX NN WO200172827-A2.
XX PD 04-OCT-2001.
XX PF 23-MAR-2001; 2001WO-US09470.
XX PR 24-MAR-2000; 2000US-191863P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Glucksmann MA;
XX DR WPI: 2001-626254/72.
XX PT New polypeptide for preventing or treating disorders associated with
XX PT cellular adhesion, proliferation or differentiation, comprises
XX PT polypeptide 33395, a member of the leucine rich repeat protein family
XX
XX Disclosure: Fig 3A-B; 133pp; English.
CC The invention provides an isolated nucleic acid encoding a polypeptide
CC of the leucine rich repeat (LRR) family, designated 33395. The 33395
CC polypeptide can be expressed by standard recombinant methodology. The
CC 33395 polynucleotides and polypeptide can be used to prevent or treat
CC disorders associated with 33395 expression, for example those involving
CC aberrant cellular adhesion, proliferation or differentiation. Specific
CC examples include melioidosis, juvenile AIDS, diabetes mellitus,
CC rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis,
CC Crohn's disease, ulcerative colitis, asthma, anemia, and chronic active
CC hepatitis. Sequences AAG65806-814 represent consensus amino acid
CC sequences derived from a hidden Markov model corresponding to the LRR
CC domains of the the human 33395 polypeptide.
XX
XX Sequence 55 AA:
SQ
Query Match 25 % Score 73.5 DB 22 Length 55;
Best Local Similarity 30.8% Pred. NO.0.032;
Matches 16; Conservative 9; Mismatches 22; Indels 5; Gaps 1
QY 1 NPMACDCRARPIMAFORARVSS-----SDVTCAPPEROGRLRALREADF 47
   ||| | | | : | | | | | | | | | | | | | | | | | | |
Db 1 NPFNCDCELRWLLRWLEAONNEALDDPVSSLCGASPESLRGPLLLLPSEF 52
RESULT 12
AAU64229
AAU64229 standard; Protein: 75 AA.
AAU64229;
AAU64229;
XX 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #25125.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitits; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteophtic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.

```

```

XX SHeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhalla A;
PI L' Maisonneuve J, Zhang Y, Jen S, Carter D;
DR WPI: 2001-616774/71.
XX N-PSDB: AAS59640.
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX Example 1: SEQ ID NO 25424: 1069pp; English.
XX Sequences AAU39105-AAUG68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis), acne,
CC pustulosis, hyperostosis and osteomyelitis), warts and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 75 AA:
SQ
OY Query Match 21.8%; Score 62; DB 22; Length 75;
Best Local Similarity 34.0%; Pred. No. 1.4;
Matches 16; Conservative 4; Mismatches 27; Indels 0; Gaps 0;
Db 3 WACDCRARPPLWAMFQARVSSDYVCATPPEHOGDRLALREADFOA 49
| | | | | | | | | | | | | | | | | | | |
| 8 WTLPKRDLSLTRSFRAETRTLRAPCOTNPKRSKDKRRKRICRYQA 54
RESULT 13
AAU65387
ID AAU65387 standard; Protein: 99 AA.
XX AAU65387;
AC
XX 27-FEB-2002 (first entry)
DT
XX Propionibacterium acnes immunogenic protein #26283.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX warts; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX WO200181581-A2.
PN
XX 01-NOV-2001.
PD
XX 20-APR-2001; 2001WO-US12865.
PP
XX 21-APR-2000; 2000US-199047P.
XX 02-JUN-2000; 2000US-208841P.
PR
XX 07-JUL-2000; 2000US-216747P.
PA (CORI-) CORIXA CORP.

```


CC nephrotropic, antitubercular, gynecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX
 SQ Sequence 59 AA:

Query Match 20.18; Score 57; DB 21; Length 59;
 Best Local Similarity 43.5%; Pred. No. 4.9;

Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 2 PWACDCRARPLMAMFORARVSS 24

1: | | | | : | | | : |
 37 PFCGCRARSRVSWFCPAMIR 59

Search completed: April 4, 2003, 08:23:48
 Job time : 44.963 secs


```

RESULT 2
US-09-764-869-976
; Sequence 976, Application US/09764869
; Patent No. US20020061521A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764,869
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 976
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (48)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (51)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (66)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-976

```

Query Match	14.4%	Score 68.5	DB 10	Length 77
Best Local Similarity	32.8%	Pred. No. 3.8		
Matches 21	Conservative 5	Mismatches 31	Indels 7	Gaps 3

```

OY 1 PRAAPRPGSGRRAGNSSNHILGV--AAAGAPADPESTLYRDLPAEDSRROGGDAPTED 58
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7 PCGPPTPRPGSGGGGGEPTFWGKKRKGAGNPPPOPTPK-APKXHGKGERGG----EF 61
OY 59 DWYG 62
    || |
Db 62 FLWG 65

```

RESULT 3
 US-09-789-561-129
 Sequence 129, Application US/09789561
 Patent No. US20020064818A1
 GENERAL INFORMATION:
 APPLICANT: Ni et al.
 TITLE OF INVENTION: 52 Human secreted proteins
 FILE REFERENCE: P2043P1
 CURRENT APPLICATION NUMBER: US/09/789, 561
 PRIOR FILING DATE: 2001-02-22
 CURRENT APPLICATION NUMBER: PCT/US00/24008
 PRIOR FILING DATE: 2000-08-31
 PRIOR APPLICATION NUMBER: 60/152,317
 PRIOR FILING DATE: 1999-09-03
 PRIOR APPLICATION NUMBER: 60/152,315
 PRIOR FILING DATE: 1999-09-03
 NUMBER OF SEQ ID NOS: 194
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 129
 LENGTH: 94
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-789-561-129

Query Match	13.7%;	Score 65;	DB 10;	Length 94;
Best Local Similarity	31.9%;	Pred. No. 10;		
Matches 23;	Conservative 6;	Mismatches 15;	Indels 28;	Gaps 4;

```

Oy      24 VAEGAPPA----DPSTLYRDLEAEDSRGQGDAETEDD-----YWGYGCGEDQGE 72
      1 : : ||| 1 1 ||| 1 : ||| 1 1 : ||: 1 1
Db      37 VSTSPPPATRCPDESELYRD-----PGADLEADQAEARGAAEGCHPGED---- 82

```

```
QY 73 QMCPGACQAPP 84
      | | : | |
Db 83 ---PWGARRGPP 91
```

```

RESULT 4
US-10-091-504-1086
: Sequence 1086, Application US/10091504
: Publication No. US2003005908A1
: GENERAL INFORMATION:
: APPLICANT: Rossen et al.,
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC000701
: CURRENT APPLICATION NUMBER: US/10/091,504
: CURRENT FILING DATE: 2002-03-07
: NUMBER OF SEQ ID NOS: 2442
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1086
:
: LENGTH: 68
:
: TYPE: FRT
: ORGANISM: Homo sapiens
: US-10-091-504-1086

```

Query Match	13.5%;	Score 64;	DB 9;	Length 68;
Best Local Similarity	35.2%;	Pred. No. 9.2;		
Matches	19;	Conservative	1;	Mismatches 18;
				Indels 16;
				Gaps 3;

```

QY      30  PRADPSTLYRLPAEDSRGQGGDAPTEDDYWGYGGEEDQGEQMGPGACQAP  83
          ||| : ||| | ||| |
Db      1  PVSPP--HEDLPA-----DKQEGTCCEGY-----RNVSGSPGAPCGCP  38

```

```

RESULT 5
US-09-764-869-1086
: Sequence 1086, Application US/09764869
: Patent No. US20020061521A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC007
: CURRENT APPLICATION NUMBER: US/09/764,869
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ. ID NOS: 2442
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1086
:
: LENGTH: 68
:
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-764-869-1086

```

Query Match 13.5%; Score 64; DB 10; Length 86;
Best Local Similarity 35.2%; Pred. No. 9.2;
Matches 19; Conservative 1; Mismatches 18; Indels 16; Gaps 3;

```
QY      30 PPADSTLYRLPAEDSRGROGGDAPEDDYWGYGGEQDRGEOMCPGAACAP 83  
          ||| | | | | | | | | | | | | | | |  
ob       1 PPVSPRP--HELDLPA-----DKOQGCTCEEG--KRWGSCPGAPCCGP 38
```

RESULT 6
US-09-867-550-1320

GENERAL INFORMATION:
 APPLICANT: Leach, Martin D.
 APPLICANT: Mehraban, Fuad,
 APPLICANT: Conley, Pamela
 APPLICANT: Law, Debbie
 APPLICANT: Topper, James
 TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells a

;; TITLE OF INVENTION: Thereby
;; FILE REFERENCE: 21402-013 (Cura-313)
;; CURRENT APPLICATION NUMBER: US/09/867,550
;; CURRENT FILING DATE: 2001-09-20
;; PRIOR APPLICATION NUMBER: USSN 60/208,427
;; PRIOR FILING DATE: 2000-05-30
;; NUMBER OF SEQ ID NOS: 2125
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 1320
;; LENGTH: 90
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (68)
;; OTHER INFORMATION: wherein Xaa may be any one of Ala or Arg or Gln or Glu or Gly or
;; OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
;; NAME/KEY: VARIANT
;; LOCATION: (72)
;; OTHER INFORMATION: wherein Xaa may be any one of Ala or Arg or Gln or Glu or Gly or
;; OTHER INFORMATION: Leu or Lys or Pro or Ser or Thr or Val
;; US-09-867-550-1320

Query Match 12.0%; Score 57; DB 10; Length 90;
Best Local Similarity 27.2%; Pred No. 62;
Matches 25; Conservative 5; Mismatches 38; Indels 24; Gaps 4;

OY 2 PAAPRPGSRARNGSSNHLGYVAEAGAPADPSTLYRDLPAEDSRGQDAPTEDDYW 61
DB 3 PPEWRSGRGR-RAGTSGSTRAATGAGAGAAAGCRF---PA-----GGAGLGPRRA 50
OY 62 GGYGGEORGE-----QMGPGAAQ 81
DB 51 GGWTGDSRGHVAVENAXLGXGEGCGSGCR 82

RESULT 7
US-09-867-550-2044
; Sequence 2044, Application US/09867550
; Patent No. US20030082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Foad.
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1e1 Polynucleotides from Atherogenic Cells and
; TITLE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2044
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: wherein Xaa may be any one of Arg or His or Leu or Pro
; US-09-867-550-2044

Query Match 11.9%; Score 56.5; DB 10; Length 98;
Best Local Similarity 31.5%; Pred. No. 77;
Matches 23; Conservative 8; Mismatches 37; Indels 5; Gaps 2;
OY 2 PAAPRPGSRARNGSSNHLGYVAEAGAPADPSTLYRDLPAEDSRG--RGGGAPTEDD 59
DB 28 PAWTRPALPLCLSLARSMAMOMRA---PMSDDLARDGSRSLRGRNRNGGGGCRDD 84

OY 60 YMGYGEORGE 72
DB 85 DDGGGAGAVRTGE 97
RESULT 8
US-09-864-761-39016
; Sequence 39016, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39016
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003668.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 16
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 14
; OTHER INFORMATION: SWISSPROT HIT: P14378, EVALUATE 5.90e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE278817.1, EVALUATE 5.50e+00

ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-31

Query Match 11.6%; Score 55; DB 10; Length 94;
Best Local Similarity 28.1%; Pred. No. 1e+02;
Matches 16; Conservative 5; Mismatches 12; Indels 24; Gaps 2;

QY 25 AEAGAPPADPSTLYRDLPAEDSRGCGDAPTEDDYWGCGEDRGOMCPGACQ 81
DB 19 AQAAGLEAEHQAIIRDV-----LTASDFWGGAG-----SAACQ 51

RESULT 12
US-09-023-588-21
Sequence 21, Application US/09023588
Patent No. US20020081579A1
GENERAL INFORMATION:
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark R.
TITLE OF INVENTION: METHOD FOR THE ISOLATION OF NOVEL ANTIGENS
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,588
FILING DATE: 14-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.445
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-023-588-21

Query Match 11.6%; Score 55; DB 10; Length 94;
Best Local Similarity 28.1%; Pred. No. 1e+02;

Matches 16; Conservative 5; Mismatches 12; Indels 24; Gaps 2;
QY 25 AEAGAPPADPSTLYRDLPAEDSRGCGDAPTEDDYWGCGEDRGOMCPGACQ 81
DB 19 AQAAGLEAEHQAIIRDV-----LTASDFWGGAG-----SAACQ 51

RESULT 13
US-09-023-588-23
Sequence 23, Application US/09023588
Patent No. US20020081579A1
GENERAL INFORMATION:
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark R.
TITLE OF INVENTION: METHOD FOR THE ISOLATION OF NOVEL ANTIGENS
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,588
FILING DATE: 14-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.445
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-023-588-23

Query Match 11.6%; Score 55; DB 10; Length 94;
Best Local Similarity 28.1%; Pred. No. 1e+02;
Matches 16; Conservative 5; Mismatches 12; Indels 24; Gaps 2;

QY 25 AEAGAPPADPSTLYRDLPAEDSRGCGDAPTEDDYWGCGEDRGOMCPGACQ 81
DB 19 AQAAGLEAEHQAIIRDV-----LTASDFWGGAG-----SAACQ 51

RESULT 14
US-09-023-588-31
Sequence 31, Application US/09023588
Patent No. US20020081579A1
GENERAL INFORMATION:
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark R.
TITLE OF INVENTION: METHOD FOR THE ISOLATION OF NOVEL ANTIGENS
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,588
 FILING DATE: 14-FEB-1998
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Makl, David J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 210121.445
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 31:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 94 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Mycobacterium tuberculosis
 US-09-023-588-31

[illegible]

```

: RESULT 15 US-09-793-306-21
: Sequence 21, Application US/09793306
: Patent No. US2002098200A1
: GENERAL INFORMATION:
: APPLICANT: Campos-Neto, Antonio
: APPLICANT: Skelky, Yasir
: APPLICANT: Owendale, Pamela
: APPLICANT: Jen, Shyian
: APPLICANT: Lodes, Michael
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compounds and Methods for Diagnosis and Immunotherapy
: TITLE OF INVENTION: of Tuberculosis
: FILE REFERENCE: 014058-00874005
: CURRENT APPLICATION NUMBER: US/09/793,306
: CURRENT FILING DATE: 2001-02-26
: PRIOR APPLICATION NUMBER: US 60/185,037
: PRIOR FILING DATE: 2000-02-25
: PRIOR APPLICATION NUMBER: US 60/223,828
: PRIOR FILING DATE: 2000-08-08
: NUMBER OF SEQ ID NOS: 164
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 21
: LENGTH: 94
: TYPE: PR1
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: OTHER INFORMATION: Tb398 ORF-2
US-09-793-306-21

```

Query Match	11.6%	Score 55;	DB 10;	Length 94;
Best Local Similarity	28.1%	Pred. No. 1e+02;		
Matches	16;	Conservative	5;	Mismatches 12;
				Indels 24;
				Gaps 2

OY
25 AEGAPPEADESTLYRDLPAEDSRGOGDAPTEDDYWGVIYGCEQDRGEQCPACAG 81
| : || | : | : | : | : |||| |
Db 19 AQAGLEAEHQATIRDV-----LTASDFWGGAG-----SAAQC 51

Search completed: April 4, 2003, 08:31:57
Job time : 146.444 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 22.6667 Seconds
(without alignments)
110.336 Million cell updates/sec

Title: US-09-972-546-2_COPY_311_395

Perfect score: 475
Sequence: 1 PPAAPTRPGSRKRGSSNNH.....GDDRGEGMCPGACQAPPD 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters: 23178

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59.5	12.5	87	4	US-09-314-268-138 Sequence 138, App
2	54	11.4	59	3	US-08-787-739-50 Sequence 50, App
3	54	11.4	59	4	US-09-178-115-50 Sequence 50, App
4	54	11.4	59	4	US-09-177-776-50 Sequence 50, App
5	54	11.4	98	2	US-08-481-658B-50 Sequence 50, App
6	54	11.4	98	2	US-08-477-504A-50 Sequence 50, App
7	54	11.4	98	2	US-08-486-756A-50 Sequence 50, App
8	54	11.4	98	2	US-08-485-862B-50 Sequence 50, App
9	54	11.4	98	3	US-08-487-077A-50 Sequence 50, App
10	54	11.4	98	3	US-08-485-863A-50 Sequence 50, App
11	54	11.4	98	4	US-08-485-049D-50 Sequence 50, App
12	53.5	11.3	72	1	US-07-609-716-113 Sequence 113, App
13	53.5	11.3	72	4	US-08-475-411A-113 Sequence 113, App
14	53.5	11.3	72	4	US-08-478-029A-113 Sequence 113, App
15	53	11.2	69	1	US-07-609-716-112 Sequence 112, App
16	53	11.2	69	4	US-08-475-411A-112 Sequence 112, App
17	53	11.2	69	4	US-08-478-029A-112 Sequence 112, App
18	51	10.7	86	3	US-08-894-017-5 Sequence 5, App
19	50.5	10.6	46	4	US-09-297-981-20 Sequence 20, App
20	50.5	10.6	83	1	US-07-609-716-115 Sequence 115, App
21	50.5	10.6	83	4	US-08-475-411A-115 Sequence 115, App
22	50.5	10.6	83	4	US-08-478-029A-115 Sequence 115, App
23	50.5	10.6	85	4	US-09-372-422A-36 Sequence 36, App
24	50	10.5	80	1	US-08-832-883-5 Sequence 5, App
25	50	10.5	80	2	US-08-832-877-5 Sequence 5, App
26	49.5	10.4	62	1	US-08-616-732A-24 Sequence 24, App
27	49.5	10.4	62	4	US-09-037-742B-24 Sequence 24, App

28	49.5	10.4	86	2	US-08-343-443B-7	Sequence 7, App
29	48.5	10.2	60	4	US-09-336-536-13	Sequence 13, App
30	48.5	10.2	82	1	US-07-609-716-114	Sequence 114, App
31	48.5	10.2	98	2	US-08-211-202-111	Sequence 111, App
32	48.5	10.2	98	2	US-08-665-202-111	Sequence 38, App
33	48	10.1	56	3	US-08-897-924A-2	Sequence 2, App
34	48	10.1	73	4	US-09-220-528-50	Sequence 50, App
35	48	10.1	68	4	US-09-564-805-213	Sequence 213, App
36	47.5	10.0	53	4	US-08-314-268-159	Sequence 159, App
37	47	9.9	50	3	US-08-897-924A-10	Sequence 10, App
38	47	9.9	56	2	US-08-850-910A-24	Sequence 24, App
39	47	9.9	58	3	US-08-897-924A-9	Sequence 9, App
40	47	9.9	67	3	US-08-897-924A-8	Sequence 8, App
41	47	9.9	70	4	US-09-314-268-114	Sequence 114, App
42	47	9.9	85	4	US-08-475-411A-114	Sequence 114, App
43	47	9.9	85	4	US-08-478-029A-114	Sequence 114, App
44	46.5	9.8	76	2	US-08-284-391B-53	Sequence 53, App
45	46.5	9.8	76	4	US-09-218-950-53	Sequence 53, App

ALIGNMENTS

```

RESULT 1
US-09-314-268-138
; Sequence 138, Application US/09314268
; Patent No. 6346377
; GENERAL INFORMATION:
; APPLICANT: Doctbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314,268
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314,268
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 138
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Human papillomavirus type 47
US-09-314-268-138

Query Match      12.5%; Score 59.5; DB 4; Length 87;
Best Local Similarity 34.5%; Pred. No. 14;
Matches 20; Conservative 1; Mismatches 24; Indels 13; Gaps 2;

QY 1 PPAAPTRPGSRKRGSSNNHLYGVAEAGAPPADPSILYRDPAEDSRGOGDAPTED 58
DB 29 PPPPTGPPGLR-----RSTRVLVLPVGQGPDP-----DLPAPVGEVGEHPOGKD 73

RESULT 2
US-08-787-739-50
; Sequence 50, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Sylvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/787,739
        FILING DATE: 24-JAN-1997
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/485,049
          FILING DATE: 07-JUN-1995
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/486,756
          FILING DATE: 07-JUN-1995
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/477,504
          FILING DATE: 07-JUN-1995
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/481,658
          FILING DATE: 07-JUN-1995
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/485,862
          FILING DATE: 07-JUN-1995
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/485,863
          FILING DATE: 07-JUN-1995
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/487,077
          FILING DATE: 07-JUN-1995
        ATTORNEY/AGENT INFORMATION:
          NAME: Lauder, Leona L.
          REGISTRATION NUMBER: 30,863
          REFERENCE/DOCKET NUMBER: D-0021.4
          TELECOMMUNICATION INFORMATION:
            TELEPHONE: 415-981-2034
            TELEFAX: 415-981-0332
        INFORMATION FOR SEQ ID NO: 50:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 59 amino acids
            TYPE: amino acid
            TOPOLOGY: linear
          MOLECULE TYPE: protein
          DESCRIPTION: MN protein proteoglycan-like region
    US-08-787-739-50

Query Match . 11.4%; Score 54; DB 3; Length 59;
Best Local Similarity 39.5%; Pred. No. 34;
Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2

Oy 40 DIPADSRKGGDAPTDDYIGVGEDDQGEOMCPG 77
|||:: |::| |::| |::| |
12 DLPSEE-----DSPRED--PPGEDLDPGEDLPG 39

RESULT 3
US-09-178-115-50
Sequence 50, Application US/09178115
Patent No. 6297041
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
FILE REFERENCE: D-0021.5A
CURRENT APPLICATION NUMBER: US/09/178,115
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 09/177,776
EARLIER FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 08/787,739
EARLIER FILING DATE: 1997-01-24
EARLIER APPLICATION NUMBER: 08/485,049
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/486,756
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/477,504
EARLIER FILING DATE: 1995-06-07

```

```

? EARLIER APPLICATION NUMBER: 08/481,658
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/485,862
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/485,863
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/487,077
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/260,190
? EARLIER FILING DATE: 1994-06-15
? EARLIER APPLICATION NUMBER: 08/177,093
? EARLIER FILING DATE: 1993-12-30
? EARLIER APPLICATION NUMBER: 07/964,589
? EARLIER FILING DATE: 1992-10-21
? EARLIER APPLICATION NUMBER: PV-709-92
? EARLIER FILING DATE: 1992-03-11
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 50
? LENGTH: 59
? TYPE: PRT
? ORGANISM: HUMAN
US-09-178-115-50

Query Match          11.4%  Score 54;  DB 4;  Length 59;
Best Local Similarity 39.5%  Pred. No. 34;
Matches 15;  Conservative 5;  Mismatches 8;  Indels 10;  Gaps 2.

OY      40  DLPADSRGROGADPTDDYDYGXGGEQDRCQMCPG 77
         |||::: 1::1::1 11 11::11
Db       12  DLPSEE-----DSPREDD--PGEEDLPGEEDLPG 39

RESULT 4
US-09-177-776-50
? Sequence 50, Application US/09177776A
? Patent No. 6297051
? GENERAL INFORMATION:
? APPLICANT: Zavada, Jan
? APPLICANT: Pastorekova, Silvia
? APPLICANT: Pastorek, Jaromir
? TITLE OF INVENTION: MN Gene and Protein
? FILE REFERENCE: D-0021.5A
? CURRENT APPLICATION NUMBER: US/09/177,776A
? CURRENT FILING DATE: 1998-10-23
? EARLIER APPLICATION NUMBER: 08/787,739
? EARLIER FILING DATE: 1997-01-24
? EARLIER APPLICATION NUMBER: 08/485,049
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/486,756
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/477,504
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/481,658
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/485,862
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/485,863
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/487,077
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 08/260,190
? EARLIER FILING DATE: 1994-06-15
? EARLIER APPLICATION NUMBER: 08/177,093
? EARLIER FILING DATE: 1993-12-30
? EARLIER APPLICATION NUMBER: 07/964,589
? EARLIER FILING DATE: 1992-10-21
? EARLIER APPLICATION NUMBER: PV-709-92
? EARLIER FILING DATE: 1992-03-11
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 50
? LENGTH: 59

```

TYPE: PRT
ORGANISM: HUMAN
US-09-177-776-50

Query Match 11.4%; Score 54; DB 4; Length 59;
Best Local Similarity 39.5%; Pred. No. 34;
Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2;

OY 40 DLPADSRRGGDAPTEDDYWGCGEDQGEOMCPG 77
DB 12 DLPSE-----DSPRED---PGEEDLPGEEDLP 39

RESULT 5
US-08-481-658B-50
Sequence 50 Application US/08481658B
Patent No. 5953075
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of homology to collagen alpha
DESCRIPTION: 1 chain
US-08-481-658B-50
Query Match 11.4%; Score 54; DB 2; Length 98;
Best Local Similarity 39.5%; Pred. No. 61;
Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2;
OY 40 DLPADSRRGGDAPTEDDYWGCGEDQGEOMCPG 77
DB 27 DLPSE-----DSPRED---PGEEDLPGEEDLP 54
RESULT 6
US-08-477-504A-50
Sequence 50 Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:

APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477.504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of homology to collagen alpha
DESCRIPTION: 1 chain
US-08-477-504A-50

Query Match 11.4%; Score 54; DB 2; Length 98;
Best Local Similarity 39.5%; Pred. No. 61;
Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2;
OY 40 DLPADSRRGGDAPTEDDYWGCGEDQGEOMCPG 77
DB 27 DLPSE-----DSPRED---PGEEDLPGEEDLP 54

RESULT 7
US-08-486-756A-50
Sequence 50 Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 42A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of homology to collagen alpha
DESCRIPTION: 1 chain
US-08-486-756A-50

Query Match 11.4%; Score 54; DB 2; Length 98;
Best Local Similarity 39.5%; Pred. No. 61;
Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2;

OY 40 DLPADSRRGGDAPTEDDYGGYGGEDQGEQMCPCG 77
|||:|:| 1:1:1:1 1:1:1:1
Db 27 DLPSEE-----DSPRED---PPGEEDLPGEEDLP 54

RESULT 8
US-08-485-862B-50
Sequence 50, Application US/08485862B
Patent No. 3989838
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of homology to collagen alpha
DESCRIPTION: 1 chain
US-08-485-862B-50

Query Match 11.4%; Score 54; DB 2; Length 98;
Best Local Similarity 39.5%; Pred. No. 61;
Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2;

OY 40 DLPADSRRGGDAPTEDDYGGYGGEDQGEQMCPCG 77
|||:|:| 1:1:1:1 1:1:1:1
Db 27 DLPSEE-----DSPRED---PPGEEDLPGEEDLP 54

RESULT 9
US-08-487-077A-50
Sequence 50, Application US/08487077A
Patent No. 6069242
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,077A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3H
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of homology to collagen alpha
DESCRIPTION: 1 chain
US-08-487-077A-50

Query Match 11.4%; Score 54; DB 3; Length 98;
Best Local Similarity 39.5%; Pred. No. 61;
Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2;

OY 40 DLPADSRRGGDAPTEDDYGGYGGEDQGEQMCPCG 77
|||:|:| 1:1:1:1 1:1:1:1
Db 27 DLPSEE-----DSPRED---PPGEEDLPGEEDLP 54

RESULT 10
US-08-485-863A-50


```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/08/485,049D
5 FILING DATE: 07-JUN-1995
6 CLASSIFICATION: 435
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US 08/260,190
9 FILING DATE: 15-JUN-1994
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Lauder, Leona L.
12 REGISTRATION NUMBER: 30,863
13 REFERENCE/DOCKET NUMBER: D-0021.3E
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 415-981-2034
16 TELEFAX: 415-981-0332
17 INFORMATION FOR SEQ ID NO: 50:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 98 amino acids
20 TYPE: amino acid
21 TOPOLOGY: linear
22 MOLECULE TYPE: protein
23 DESCRIPTION: Region of homology to collagen alpha
24 DISCRPTION: 1 chain
25 US-08-485-049D-50
26
27 Query Match 11.4%; Score 54; DB 4; Length 98;
28 Best Local Similarity 39.5%; Pred. No. 61;
29 Matches 15; Conservative 5; Mismatches 8; Indels 10; Gaps 2.
30
31 Qy 40 DLPAEDSRGROGDAPTEDDYWGCVGGEGQREOMCPG 77
32 ||||: |:||: ||||: ||
33 Db 27 DLPSE-----DSPRED---PGEEDLPSEDLPG 54
34
35 RESULT 12
36 US-07-609-716-113
37 ; Sequence 113, Application US/07609716
38 ; Patent No. 5514581
39 ; GENERAL INFORMATION:
40 ; APPLICANT: Ferrari, Franco A.
41 ; APPLICANT: Cappello, Joseph
42 ; TITLE OF INVENTION: Functional Recombinantly Prepared
43 ; TITLE OF INVENTION: Synthetic Protein Polymer
44 ; NUMBER OF SEQUENCES: 118
45 ; CORRESPONDENCE ADDRESS:
46 ; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
47 ; STREET: Four Embarcadero Center, Suite 3400
48 ; CITY: San Francisco
49 ; STATE: CA
50 ; COUNTRY: US
51 ; ZIP: 94111
52 ; COMPUTER READABLE FORM:
53 ; MEDIUM TYPE: Floppy disk
54 ; COMPUTER: IBM PC compatible
55 ; OPERATING SYSTEM: PC-DOS/MS-DOS
56 ; SOFTWARE: Patentin Release #1.0, Version #1.30
57 ; CURRENT APPLICATION DATA:
58 ; APPLICATION NUMBER: US/07/609,716
59 ; FILING DATE: 06-NOV-1990
60 ; CLASSIFICATION: 435
61 ; ATTORNEY/AGENT INFORMATION:
62 ; NAME: Rowland, Bertram I
63 ; REGISTRATION NUMBER: 20015
64 ; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
65 ; TELECOMMUNICATION INFORMATION:
66 ; TELEPHONE: 415-781-1989
67 ; TELEFAX: 415-398-3249
68 ; INFORMATION FOR SEQ ID NO: 113:
69 ; SEQUENCE CHARACTERISTICS:
70 ; LENGTH: 72 amino acids
71 ; TYPE: amino acid
72 ; STRANDEDNESS: single

```

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-609-716-113

Query Match
Best Local Similarity 26.1%; Score 53.5; DB 1; Length 72;
Matches 24; Conservative 4; Mismatches 19; Indels 45; Gaps 6;

1 PPAPTRPGSRARNGSSNHLGYAAGAPPADPSTLYRDLPAEDSRGROG--GDAPTED 58
14 PGAGPGPG-----PPGPG-----LP-----GPKGDRGDA----- 39

59 DYWGCGEDQGEOMCP-----GAACQAPP 84
40 ---GPKGADGSPGALCVSEPGYISRCDA GP 68

RESULT 13
US-08-475-411A-113
Sequence 113, Application US/08475411A
Patent No. 6140072

GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
TITLE OF INVENTION: Functional Recombinantly Prepared
TITLE OF INVENTION: Synthetic Protein Polymer
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,411A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/609,716
FILING DATE: 06-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/269,429
FILING DATE: 09-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-9/RFT/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-475-411A-113

Query Match
Best Local Similarity 26.1%; Score 53.5; DB 4; Length 72;
Matches 24; Conservative 4; Mismatches 19; Indels 45; Gaps 6;

Matches 24; Conservative 4; Mismatches 19; Indels 45; Gaps 6;

1 PPAPTRPGSRARNGSSNHLGYAAGAPPADPSTLYRDLPAEDSRGROG--GDAPTED 58
14 PGAGPGPG-----PPGPG-----LP-----GPKGDRGDA----- 39

59 DYWGCGEDQGEOMCP-----GAACQAPP 84
40 ---GPKGADGSPGALCVSEPGYISRCDA GP 68

RESULT 14
US-08-478-029A-113
Sequence 113, Application US/08478029A
Patent No. 6184348

GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
TITLE OF INVENTION: Functional Recombinantly Prepared
TITLE OF INVENTION: Synthetic Protein Polymer
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,029A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/609,716
FILING DATE: 06-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/269,429
FILING DATE: 09-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-8/RFT/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-478-029A-113

Query Match
Best Local Similarity 26.1%; Score 53.5; DB 4; Length 72;
Matches 24; Conservative 4; Mismatches 19; Indels 45; Gaps 6;

1 PPAPTRPGSRARNGSSNHLGYAAGAPPADPSTLYRDLPAEDSRGROG--GDAPTED 58
14 PGAGPGPG-----PPGPG-----LP-----GPKGDRGDA----- 39

OY 59 DYWGCGEDRGGEOMCP-----GACQAPP 84
 DB 40 ----GPKGADGSPGALCVSEPGYIGSRCDAGP 68

RESULT 15

US-07-609-716-112

; Sequence 112; Application US/07609716
 ; Patent No. 5514581

GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.
 APPLICANT: Cappello, Joseph
 TITLE OF INVENTION: Functional Recombinantly Prepared
 TITLE OF INVENTION: Synthetic Protein Polymer
 NUMBER OF SEQUENCES: 118

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
 STREET: Four Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: CA
 COUNTRY: US
 ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/609,716
 FILING DATE: 06-NOV-1990
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rowland, Berttram I
 REGISTRATION NUMBER: 20015
 REFERENCE/DOCKET NUMBER: A-55186-3/BIR
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-781-1989
 TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 112:

SEQUENCE CHARACTERISTICS:
 LENGTH: 69 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-07-609-716-112

Query Match

11.2% Score 53; DB 1; Length 69;
 26.7% Identity; 4; Mismatches 18; Indels 44; Gaps 6;

atches 24; Conservative 4; Mismatches 18; Indels 44; Gaps 6;

OY 1 PPAALTRPGSRARGNSSNHLVGAAGAPADPSTLYRDLPAEDSRGROG--GDAPTED 58
 DB 14 PCGAPGPPG-----PPGPPG-----LP-----GPKGDRGDA----- 39
 OY 59 DYWGCGEDRGGEOMCP-----GACQAPP 84
 DB 40 ----GPKGADGSPGALCVSEPGYIGSRCDAGP 65

Search completed: April 4, 2003, 08:26:07
 Job time: 23.6667 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 66.1111 Seconds
(without alignments)
264.918 Million cell updates/sec

Title: US-09-972-546-2_COPY_311_395

Perfect score: 475
Sequence: 1 PPAAPTRPGSRANRGNSSNNH.....GEDRGRCMCGAACAPPD 85

Scoring table:
BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

number of hits satisfying chosen parameters: 101114

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62.5	13.2	70	12	09Q0D8 herpes simp
2	62.5	13.2	70	12	09Q0D7 herpes simp
3	62.5	13.2	70	12	09Q0D6 herpes simp
4	62.5	13.2	70	12	09Q0D5 herpes simp
5	61	12.8	70	12	09Q0D0 herpes simp
6	61	12.8	70	12	09Q0C9 herpes simp
7	61	12.8	92	3	096T52 herpes simp
8	60.5	12.7	93	4	015215 schizoply11
9	59.5	12.5	74	4	015905 homo sapien
10	58.5	12.3	73	4	005200 homo sapien
11	58	12.2	68	4	09UD2 homo sapien
12	57	12.0	89	4	016403 homo sapien
13	56.5	11.9	60	16	053824 mycobacteri
14	56.5	11.9	70	12	09Q0D4 herpes simp
15	56.5	11.9	70	12	09Q0D3 herpes simp
16	56.5	11.9	70	12	09Q0D2 herpes simp

17	56.5	11.9	80	12	070795	070795	tt virus. s
18	56.5	11.9	80	12	070799	070799	tt virus. s
19	56	11.8	90	4	09BY99	09BY99	homo sapien
20	55.5	11.7	70	12	09Q0D1	09Q0D1	herpes simp
21	55	11.6	68	16	0984J4	0984J4	rhizobium 1
22	55	11.6	91	10	09EP40	09EP40	oryza sativ
23	54.5	11.5	67	4	09UH23	09UH23	homo sapien
24	54	11.4	78	10	094HX9	094HX9	oryza sativ
25	54	11.4	87	5	09VQ08	09VQ08	oryza sativ
26	53.5	11.3	53	12	065541	065541	bovine herp
27	53.5	11.3	95	5	09VRK7	09VRK7	dirosophila
28	53.5	11.3	95	5	08WT20	08WT20	ostertagia
29	53	11.2	65	16	082RK6	082RK6	salmonella
30	53	11.2	70	4	095527	095527	homo sapien
31	53	11.2	72	4	015203	015203	homo sapien
32	53	11.2	80	10	09FUP5	09FUP5	phaeocystis
33	53	11.2	89	13	091413	091413	notophthal
34	53	11.2	90	10	09ATH2	09ATH2	oryza sativ
35	53	11.2	98	13	09PVQ3	09PVQ3	oryza sativ
36	52.5	11.1	56	6	09XSA9	09XSA9	bos taurus
37	52.5	11.1	57	16	09K3S2	09K3S2	streptococ
38	52.5	11.1	90	16	P73629	P73629	synecocyst
39	52.5	11.1	99	5	09VPS9	09VPS9	dirosophila
40	52	10.9	87	5	024703	024703	dirosophila
41	52	10.9	95	2	09EUN9	09EUN9	corynebacte
42	51.5	10.8	56	16	092KB9	092KB9	rhizobium m
43	51.5	10.8	69	12	09PZ57	09PZ57	simian herp
44	51.5	10.8	90	5	018562	018562	strongyloid
45	51.5	10.8	92	12	067789	067789	human adeno

ALIGNMENTS

RESULT 1

ID	09Q0D8	PRELIMINARY:	PRT:	70 AA.
AC	09Q0D8;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Glycoprotein G-2 (Fragment).			
CN	GG-2.			
OS	Herpes simplex virus (type 2).			
OC	Viruses; dsDNA viruses, no RNA stage: Herpesviridae;			
OC	Alphaherpesvirinae; Simplexvirus.			
OX	NCBI_TaxID=10310;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-VI-2944;			
RX	MEDLINE=20553357; PubMed=11101589;			
RA	Liljeqvist J.A., Svernerholm B., Bergstrom T.;			
RT	"Conservation of type-specific B-cell epitopes of glycoprotein G in			
RT	clinical herpes simplex virus type 2 isolates.";			
RL	J. Clin. Microbiol. 38:4517-4522(2000).			
DR	EMBL: AJ270565; CAB58428.1;			
DR	InterPro: IPR003363; Herpes_gg.			
DR	Pfam: PF02400; Herpes_gg; 1.			
FT	NON_TER	1		
FT	NON_TER	70		
SO	SEQUENCE	70 AA;	7130 MW;	051CA419B1A5440F CRC64;
Query Match				
Best Local Similarity 13.2%; Score 62.5; DB 12; Length 70;				
Matches 22; Conservative 10; Mismatches 35; Indels 15; Gaps 4;				
OY	3 AAP-TRGSRARNGSSNHLGVACAPAPDPSTLYRDLPAEDSRGSGDAPETDDYW 61			
DB	3 ATPGTRGARTPPDPKTYHGPAD--APPSPAP-----PPENRGR-----PEFFE 48			
OY	62 GYGCGEDRGRCMCGAACAP 83			
DB	49 GAGGDEPPDDSDATGIAFRP 70			

```
RESULT 2
0900D7 ID 0900D7 PRELIMINARY; PRT: 70 AA.
AC 0900D7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycoprotein G-2 (Fragment).
GN GG-2.
OS Herpes simplex virus (type 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VI-449;
MEDLINE=20553357; PubMed=11101589;
Liljegyist J.A., Svennerholm B., Bergstrom T.;
"Conservation of type-specific B-cell epitopes of glycoprotein G in
clinical herpes simplex virus type 2 isolates.";
J. Clin. Microbiol. 38:4517-4522(2000).
DR EMBL; AJ270566; CAB58429.1; -
DR InterPro; IPR003363; Herpes_gg.
DR Pfam; PF02400; Herpes_gg; 1.
FT NON_TER 1
FT SEQUENCE 70 AA; 7130 MW; 051CA419B1A5440F CRC64;

Query Match 13.2%; Score 62.5; DB 12; Length 70;
Best Local Similarity 26.8%; Pred. No. 31;
Matches 22; Conservative 10; Mismatches 35; Indels 15; Gaps 4;

OY 3 AAP-TRPGSRARGNSSNHLGVAEAGAPADPSTLYRDLPEDSRGQGGAPTEDDY 61
DB 3 ATPGTRGTRARPTPTDPKTHPHGPAD--APGSPAD-----PPEHRGR-----PEEFE 48
OY 62 GYGGEQDQGEQMCPCGAACQAP 83
DB 49 GAGDGEPPEDDDSATGLAFRTP 70

RESULT 3
0900D6 ID 0900D6 PRELIMINARY; PRT: 70 AA.
AC 0900D6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycoprotein G-2 (Fragment).
GN GG-2.
OS Herpes simplex virus (type 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VI-448;
MEDLINE=20553357; PubMed=11101589;
Liljegyist J.A., Svennerholm B., Bergstrom T.;
"Conservation of type-specific B-cell epitopes of glycoprotein G in
clinical herpes simplex virus type 2 isolates.";
J. Clin. Microbiol. 38:4517-4522(2000).
DR EMBL; AJ270567; CAB58430.1; -
DR InterPro; IPR003363; Herpes_gg.
DR Pfam; PF02400; Herpes_gg; 1.
FT NON_TER 1
FT SEQUENCE 70 AA; 7130 MW; 051CA419B1A5440F CRC64;

Query Match 13.2%; Score 62.5; DB 12; Length 70;
Best Local Similarity 26.8%; Pred. No. 31;
Matches 22; Conservative 10; Mismatches 35; Indels 15; Gaps 4;
```

```
Matches 22; Conservative 10; Mismatches 35; Indels 15; Gaps 4;

OY 3 AAP-TRPGSRARGNSSNHLGVAEAGAPADPSTLYRDLPEDSRGQGGAPTEDDY 61
DB 3 ATPGTRGTRARPTPTDPKTHPHGPAD--APGSPAD-----PPEHRGR-----PEEFE 48
OY 62 GYGGEQDQGEQMCPCGAACQAP 83
DB 49 GAGDGEPPEDDDSATGLAFRTP 70

RESULT 4
0900D5 ID 0900D5 PRELIMINARY; PRT: 70 AA.
AC 0900D5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycoprotein G-2 (Fragment).
GN GG-2.
OS Herpes simplex virus (type 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VI-429;
MEDLINE=20553357; PubMed=11101589;
Liljegyist J.A., Svennerholm B., Bergstrom T.;
"Conservation of type-specific B-cell epitopes of glycoprotein G in
clinical herpes simplex virus type 2 isolates.";
J. Clin. Microbiol. 38:4517-4522(2000).
DR EMBL; AJ270568; CAB58431.1; -
DR InterPro; IPR003363; Herpes_gg.
DR Pfam; PF02400; Herpes_gg; 1.
FT NON_TER 1
FT SEQUENCE 70 AA; 7130 MW; 051CA419B1A5440F CRC64;

Query Match 13.2%; Score 62.5; DB 12; Length 70;
Best Local Similarity 26.8%; Pred. No. 31;
Matches 22; Conservative 10; Mismatches 35; Indels 15; Gaps 4;

OY 3 AAP-TRPGSRARGNSSNHLGVAEAGAPADPSTLYRDLPEDSRGQGGAPTEDDY 61
DB 3 ATPGTRGTRARPTPTDPKTHPHGPAD--APGSPAD-----PPEHRGR-----PEEFE 48
OY 62 GYGGEQDQGEQMCPCGAACQAP 83
DB 49 GAGDGEPPEDDDSATGLAFRTP 70

RESULT 5
0900D0 ID 0900D0 PRELIMINARY; PRT: 70 AA.
AC 0900D0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycoprotein G-2 (Fragment).
GN GG-2.
OS Herpes simplex virus (type 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VI-1070;
MEDLINE=20553357; PubMed=11101589;
Liljegyist J.A., Svennerholm B., Bergstrom T.;
"Conservation of type-specific B-cell epitopes of glycoprotein G in
clinical herpes simplex virus type 2 isolates.";
J. Clin. Microbiol. 38:4517-4522(2000).
```

DR EMBL: AJ270573; CAB58436.1; -
DR InterPro: IPR003363; Herpes_gg.
DR Pfam: PF02400; Herpes_gg; 1.
FT NON_TER 1
FT NON_TER 70
SQ SEQUENCE 70 AA: 7091 MW: 45F05E0511E72338 CRC64;

Query Match 12.8%; Score 61; DB 12; Length 70;
Best Local Similarity 32.7%; Pred. No. 43;
Matches 18; Conservative 7; Mismatches 28; Indels 2; Gaps 2;

OY 6 TRGSGARCNSSNHLGYVAEA-GAPPADPSTLYRDLPAEDSRGCGDAPTEDD 59
DB 7 TGTGARTPTDPTKTHPGTDAPEGSPAPSPRRHGRG-EEFEGAGDGEPEDDDD 60

RESULT 6
O900C9 PRELIMINARY: PRT: 70 AA.
O900C9.

01-MAY-2000 (TREMBlrel. 13, Created)
01-MAY-2000 (TREMBlrel. 13, Last sequence update)
01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Glycoprotein G-2 (Fragment).
GN G-2.

OS Herpes simplex virus (type 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VI-1915;
RX MEDLINE=20553357; PubMed=11101589;

RA Lillqvist J.A., Sjönerholm B., Bergström T.;
RT "Conservation of type-specific B-cell epitopes of glycoprotein G in
clinical herpes simplex virus type 2 isolates."
RL J. Clin. Microbiol. 38:4517-4522(2000).
DR EMBL: AJ270574; CAB58437.1; -

DR InterPro: IPR003363; Herpes_gg.
DR Pfam: PF02400; Herpes_gg; 1.
FT NON_TER 1
FT NON_TER 70
SQ SEQUENCE 70 AA: 7091 MW: 45F05E0511E72338 CRC64;

Query Match 12.8%; Score 61; DB 12; Length 70;
Best Local Similarity 32.7%; Pred. No. 43;
Matches 18; Conservative 7; Mismatches 28; Indels 2; Gaps 2;

OY 6 TRGSGARCNSSNHLGYVAEA-GAPPADPSTLYRDLPAEDSRGCGDAPTEDD 59
DB 7 TGTGARTPTDPTKTHPGTDAPEGSPAPSPRRHGRG-EEFEGAGDGEPEDDDD 60

RESULT 7
O96TS2 PRELIMINARY: PRT: 92 AA.
O96TS2.

01-DEC-2001 (TREMBlrel. 19, Created)
01-DEC-2001 (TREMBlrel. 19, Last sequence update)
01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Peptide mating pheromone precursor Bbp2-5.
GN Bbp2-5.

OS Schizophyllum commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Schizophyllaceae; Schizophyllum.
OX NCBI_TaxID=5334;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4-8;
RX "Changes in mate recognition through alterations of pheromones and
RT receptors in the multisexual mushroom fungus Schizophyllum commune."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF378294; AAK58070.1; -
DR InterPro: IPR001230; Prey1 site.
DR PROSITE: PS00294; PREY1LATION; UNKNOWN_1.
SQ SEQUENCE 92 AA: 9594 MW: A3CB0180AB379C03 CRC64;

Query Match 12.8%; Score 61; DB 3; Length 92;
Best Local Similarity 32.3%; Pred. No. 59;
Matches 20; Conservative 4; Mismatches 22; Indels 16; Gaps 2;

OY 3 AAPTREGSARCNSSNHLGYVAEAGAPPADPSTLYRDLPAEDSRGCGDAPTEDDYWG 62
DB 43 STTRPAS-----ATPRAPDADLLRLADARSAIARPPQDADPP--DGYFA 86

OY 63 GY 64
DB 87 GY 88

RESULT 8
O15215 PRELIMINARY: PRT: 93 AA.
O15215.

01-NOV-1996 (TREMBlrel. 01, Created)
01-NOV-1996 (TREMBlrel. 01, Last sequence update)
01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Salivary proline-rich protein 1 (Fragment).
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298176; PubMed=5089212;
RA Azen E.A., Lyons K.M., McGonigal T., Barrett N.L., Clements L.S.,
RA Maeda N., Vanin E.F., Carlson D.M., Smiths O.;
RT "Clones from the human gene complex coding for salivary proline-rich
RT proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 81:5561-5565(1984).

DR EMBL: K02576; AAA36503.1; -
FT NON_TER 1
FT NON_TER 93
SQ SEQUENCE 93 AA: 9170 MW: 04AD5F0797E31867 CRC64;

Query Match 12.7%; Score 60.5; DB 4; Length 93;
Best Local Similarity 27.4%; Pred. No. 67;
Matches 23; Conservative 6; Mismatches 34; Indels 21; Gaps 4;

OY 1 PPAPTREGSARCNSSNHLGYVAEAGAPPADPSTLYRDLPAEDSRGCGDAPTEDDY 60
DB 18 PPGKPPQPPAPGCGCKSQS-----ARAPPKP-----QSPQEGNNPQGPPIPA----- 61

OY 61 WGVGGEEDGREGWCPGACGAPR 84
DB 62 ----GNNPQ-PQAPRAGOPGPP 80

RESULT 9
O15905 PRELIMINARY: PRT: 74 AA.
O15905.

01-NOV-1996 (TREMBlrel. 01, Created)
01-NOV-1996 (TREMBlrel. 01, Last sequence update)
01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE YB-1 protein.
GN YB-1.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96226173; PubMed=8657568;
RA Makino Y., Ohga T., Toh S., Kolike K., Okumura K., Wada M., Kuwano M.,
RA Kohno K.;

[illegible]

RA	Thalmann E., Hassfeld W., Bartla A., Smolen J.S.!
RT	"Purification and partial sequencing of the nuclear autoantigen RA33
RT	shows that it is indistinguishable from the A2 protein of the
RT	heterogeneous nuclear ribonucleoprotein complex."
RL	J. Clin. Invest. 90:1061-1066(1992).
DR	HSSP: P09651; HA1.
DR	Interpro: IPR000504; RNA_rec.mot.
DR	PROSITE; PS00030; RRM_RNP_1; UNKNOWNM_1.
FT	NON_TER 1 1
FT	NON_CONS 8 9
FT	NON_CONS 23 24
FT	NON_CONS 48 49
FT	NON_TER 68 68
SQ	SEQUENCE 68 AA; 6747 MW; F83950E4IE6C6931 CRC64;
Query Match	12.2%; Score 58; DB 4; Length 68;
Best Local Similarity	28.8%; Pred. No. 83;
Matches 15; Conservative	5; Mismatches 26; Indels 6; Gaps 1.
OY	15 NSSNNLLYVAEAGAPADPSTLYRDLPAEDSRGKGSGAAPTDEDDYWGSGYG 66 : : : : : : : : : : : : : : : : : : : DB 22 DKGGNGFGFDSRCGGGNFGPG-----PSGNFRGYGGGGGPYGNGGGTGG 67

```

RESULT 12
016403          PRELIMINARY;          PRT;          89 AA.
AC      016403;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      Type III collagen alpha 1(III) chain (Fragment).
GN      COL3A1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RP      MEDLINE=96067614; PubMed=7487954;
RA      Chiodo A.A., Sillence D.O., Cole W.G., Bateman J.F.;
RT      "Abnormal type III collagen produced by an exon-IV-skipping mutation
RT      of the COL3A1 gene in Ehlers-Danlos syndrome type IV is not
RT      incorporated into the extracellular matrix.";
RL      Biochem. J. 311:939-943(1995).
DR      EMBL; S79877; AAB35615.1;-.
DR      InterPro; IPR000087; Collagen.
DR      Pfam; PF01391; Collagen; 1.
KW      Collagen.
KW      NON TER.
SO      SEQUENCE          1          1
          89 AA; 7862 MW; 03CEA17273235A52 CRC64;

Query Match          12.0%; Score 57; DB 4; Length 89;
Best Local Similarity 28.6%; Pred. No. 1,4e+02;
Matches 26; Conservative 6; Mismatches 39; Indels 20; Gaps 6.

QY      2 PAAPRPG-----SRAGNSSNNHLYGYAAEAGCAPADPSTLYRLDLPALDSRGRGGDAPTE 57
          | 111 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4 PGRRRPGLPGAGARGNDGAR---GSDGCGPPGPGRGT-AGFP-----GSPGANGPGR 53
          | 111 | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      58 DDYWGYYGGEODRGEMCPGA-----ACQAP 84
          : 111 : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      54 IN--GSPGKGEMGPAGIGAPGLMGARGR 82
          : 111 : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
ID      053824          PRELIMINARY;          PRT;          60 AA.
AC      053824;
DT      01-JUN-1998 (TREMBLrel. 06, Created)
DT      01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT      01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE      Hypothetical protein Rv0810C.

```


CN RV0810C OR MT043.02C OR MT0831.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Filum: Actinobacteria; Actinobacteridae;
 OX Actinomycetales; Corynebacteriaceae; Mycobacterium.
 RN NCBI_TaxID=1773;
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeller K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jorgels K., Krogan A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborn J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence."
 RT Nature 393:537-544(1998).
 [2]

SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.F., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains."
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE0022004; CAAL7616.1;
 DR EMBL: AE006973; AAK45073.1;
 DR TIGR: MT0831;
 DR TubercuList: RV0810c;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 60 AA; 6900 MW; 486630323fBCC826 CRC64;

Query Match 11.9%; Score 56.5; DB 16; Length 60;
 Best Local Similarity 27.5%; Pred. No. 1e+02;
 Matches 14; Conservative 8; Mismatches 28; Indels 1; Gaps 1;
 OY 11 RARSSSSNNHLYGVAEAGAPADPSTLYRDLPAEDSRGOGGADAPTEDD 61
 Db 3 RGRKAKOTKVAELKISSPOTDFORLQRLSGTGT-DRLDGDGSPDDSD 52

RESULT 14
 ID4
 AC 0900D4 PRELIMINARY; PRT: 70 AA.
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 GN Glycoprotein G-2 (Fragment).
 CG 2.
 OS Herpes simplex virus (type 2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OX NCBI_TaxID=10310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VT-287;
 RX MEDLINE=2055357; PubMed=11101589;
 RA Liljeqvist J.A., Svennerholm B., Bergstrom T.;
 RT "Conservation of type-specific B-cell epitopes of glycoprotein G in
 clinical herpes simplex virus type 2 isolates."
 RL J. Clin. Microbiol. 38:4517-4522(2000).
 DR EMBL: AJ270569; CAB58432.1;
 DR InterPro: IPR003363; Herpes_gg.
 DR Pfam: PF02400; Herpes_gg; 1.
 FT NON_TER 1 70
 FT 70

SQ SEQUENCE 70 AA; 7029 MW; 0509E559E0E0440F CRC64;

Query Match 11.9%; Score 56.5; DB 12; Length 70;
 Best Local Similarity 32.2%; Pred. No. 1.2e+02;
 Matches 19; Conservative 8; Mismatches 29; Indels 3; Gaps 3;

OY 3 AAP-TRPGSRARSSNNHLYGVAEAGAPADPSTLYRDLPAEDSRGOGGADAPTEDD 59
 Db 3 ATPGTRCTARTPTPTDPTPHGPDADPGSPAPPPPHRGGP-QEPEGAGDGEPPDD 60

RESULT 15

ID 0900D3 PRELIMINARY; PRT: 70 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 GN Glycoprotein G-2 (Fragment).
 CG 2.

OS Herpes simplex virus (type 2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OX Alphaherpesvirinae; Simplexvirus.
 RN NCBI_TaxID=10310;
 RP [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=VT-1529;
 RX MEDLINE=2055357; PubMed=11101589;
 RA Liljeqvist J.A., Svennerholm B., Bergstrom T.;
 RT "Conservation of type-specific B-cell epitopes of glycoprotein G in
 clinical herpes simplex virus type 2 isolates."
 RL J. Clin. Microbiol. 38:4517-4522(2000).
 DR EMBL: AJ270570; CAB58433.1;
 DR InterPro: IPR003363; Herpes_gg.
 DR Pfam: PF02400; Herpes_gg; 1.
 FT NON_TER 1 70
 FT 70
 SQ SEQUENCE 70 AA; 7029 MW; 0509E559E0E0440F CRC64;

Query Match 11.9%; Score 56.5; DB 12; Length 70;
 Best Local Similarity 32.2%; Pred. No. 1.2e+02;
 Matches 19; Conservative 8; Mismatches 29; Indels 3; Gaps 3;

OY 3 AAP-TRPGSRARSSNNHLYGVAEAGAPADPSTLYRDLPAEDSRGOGGADAPTEDD 59
 Db 3 ATPGTRCTARTPTPTDPTPHGPDADPGSPAPPPPHRGGP-QEPEGAGDGEPPDD 60

Search completed: April 4, 2003, 08:28:04
 Job time: 69.1111 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 : Search time 17.6296 Seconds

(without alignments)
199.975 Million cell updates/sec

Title: US-09-972-546-2_COPY_311_395

Perfect score: 475
Sequence: 1 PPAAPRPGSRARGNSSNH.....GPDGRGOMCPGACQAAPPD 85

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 10105

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	63	13.3	85	1	MRKD_RAT
2	60.5	12.7	86	1	CA1C_BOVIN
3	56.5	11.9	85	1	Y9KD_STRLI
4	55.5	11.7	95	1	VE5B_BPVA
5	55	11.6	94	1	ES64_MYCTU
6	53	11.2	95	1	OSTR_MOUSE
7	51	10.7	94	1	ES61_MYCTU
8	50	10.5	87	1	UV15_SCHPO
9	49	10.3	94	1	ES66_MYCTU
10	48.5	10.2	93	1	NEUI_ANSAN
11	48.5	10.2	96	1	PRP5_HUMAN
12	48	10.1	96	1	GRP7_DAVCA
13	48	10.1	99	1	VE4_HPV28
14	47	9.9	96	1	GATC_DEIRA
15	46	9.7	71	1	NIFU_FRAAL
16	45.5	9.6	53	1	RUBR_CHILT
17	45.5	9.6	88	1	SEB2_YEAST
18	45	9.5	67	1	YIFL_ECOLI
19	45	9.5	70	1	DSSI_HUMAN
20	45	9.5	87	1	Y898_MYCTU
21	45	9.5	94	1	ES65_MYCTU
22	44.5	9.4	63	1	ITR4_LUPCY
23	44.5	9.4	80	1	PSAC_PINTH
24	43.5	9.2	61	1	ANK7_PINTH
25	43	9.1	72	1	ANK7_BOVIN
26	43	9.1	75	1	TRBK_AGRTO
27	43	9.1	89	1	H173_HUMAN
28	42.5	8.9	63	1	YEBW_ECOLI
29	42.5	8.9	89	1	H017_RAT
30	42.5	8.9	90	1	IT14_ARATH
31	42	8.8	100	1	MERT_STRLI
32	41.5	8.7	54	1	GALA_CANFA
33	41.5	8.7	80	1	PSAC_ARATH

34	41.5	8.7	80	1	PSAC_CYAPA
35	41.5	8.7	80	1	PSAC_GNECN
36	41.5	8.7	80	1	PSAC_MAIZE
37	41.5	8.7	80	1	PSAC_ORISA
38	41.5	8.7	80	1	PSAC_SPIOL
39	41.5	8.7	90	1	YJ93_MYCTU
40	41.5	8.7	92	1	Y5V3_CAELT
41	41.5	8.7	93	1	CD3G_BOVIN
42	41.5	8.7	96	1	ICTL_PIG
43	41.5	8.7	99	1	HG14_HUMAN
44	41	8.6	69	1	DISB_ECHCS
45	41	8.6	84	1	RS16_RALSO

ALIGNMENTS

RESULT 1	ID	MRKD_RAT	STANDARD	PRT	85 AA.
AC	P20468				
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DE	Protein Kinase C substrate 80 kDa protein (Fragments).				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE.				
RC	TISSUE=Brain;				
RX	MEDLINE=90005952; PubMed=2676596;				
RA	ERLINSKY J.D., Morris C., Perks K., Brown R., Brooks S.,				
RA	Rozengurt E.;				
RT	"Internal amino acid sequence analysis of the 80 kDa protein kinase C				
RT	substrate from rat brain: relationship to the 87 kDa substrate from				
RL	bovine brain.";				
FEBS Lett. 255:149-153(1989).					
CC	-1- FUNCTION: CELLULAR SUBSTRATE FOR PROTEIN KINASE C.				
CC	-1- SIMILARITY: TO 87 MYRISTOYLATED ALANINE-RICH C KINASE SUBSTRATE				
CC	(MARCKS).				
DR	PIR: S05681; S05681.				
KW	Phosphorylation.				
FT	NON_TER 1				
FT	NON_CONS 20				
FT	NON_CONS 49				
FT	NON_CONS 73				
FT	NON_CONS 74				
SQ	SEQUENCE 85 AA: 7593 MW: 395010786C12E268 CRC64;				
Query Match	13.3%; Score 63; DB 1; Length 85;				
Best Local Similarity	30.1%; Pred. No. 16;				
Matches	25; Conservative 6; Mismatches 30; Indels 22; Gaps 4;				
OY	3 AAPPFGSRARGNSSNLLYGAEGAPPAAPSTLYRLPAPDSGRGCGAAPPEDDYWG 62				
DB	3 AAAPRPGSRARGNSSNLLYGAEGAPPAAPSTLYRLPAPDSGRGCGAAPPEDDYWG 62				
OY	63 GYGGE-DORGEOMCPGACQAAPP 84				
DB	42 AEGAGEGKEAPGAPGATADAP 64				
RESULT 2					
ID	CA1C_BOVIN				
AC	P25508				
DT	01-MAY-1992 (Rel. 22, Created)				
DT	01-MAY-1992 (Rel. 22, Last sequence update)				
DT	01-NOV-1995 (Rel. 32, Last annotation update)				
DE	Collagen alpha 1(XII) chain (Fragments).				
CN	COL12A1.				
OS	Bos taurus (Bovine).				

```

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN
  [1]
RN MEDLINE=88255287; PubMed=1133242;
RA Dublet B., Dixon E., de Miguel E., van der Rest M.;
RT "Bovine type XII collagen: amino acid sequence of a 10 kDa pepsin
RT fragment from periodontal ligament reveals a high degree of homology
RT with the chicken alpha 1(XII) sequence.";
RL FEBS Lett. 233:177-180(1988).
CC -1- FUNCTION: TYPE XII COLLAGEN INTERACTS WITH TYPE I COLLAGEN-
CC CONTAINING FIBRILS. THE COL1 DOMAIN COULD BE ASSOCIATED WITH THE
CC SURFACE OF THE FIBRILS. AND THE COL2 AND NC3 DOMAINS MAY BE
CC LOCALIZED IN THE PERIFIBRILLAR MATRIX.
CC -1- SUBUNIT: TRIMER OF IDENTICAL CHAINS EACH CONTAINING 190 KDA OF
CC NONTRIPE-HELICAL SEQUENCES.
CC -1- PPM: THE TRIPLE-HELICAL TAIL IS STABILIZED BY DISULFIDE BONDS AT
CC EACH END.
CC -1- PPM: PROLINES AT THE THIRD POSITION OF THE TRIPLEPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- SIMILARITY: HIGH TO COLLAGEN ALPHA 1(IX).
DR PIR; S00802; S00802.
DR InterPro; IPR000087; Collagen.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cell adhesion; Collagen.
FT NON_TER 1
FT NON_CONS 15 16
FT MOD_CONS 48 49
FT MOD_RES 6 6
FT MOD_RES 9 9
FT MOD_RES 12 12
FT MOD_RES 18 18
FT MOD_RES 24 24
FT MOD_RES 27 27
FT MOD_RES 30 30
FT MOD_RES 42 42
FT MOD_RES 51 51
FT MOD_RES 54 54
FT MOD_RES 65 65
FT MOD_RES 74 74
FT MOD_RES 77 77
FT MOD_RES 80 80
FT NON_TER 86
SQ SEQUENCE 86 AA; 8062 MW; 163A35AF53EA48D CRC64;

Query Match 12.7%; Score 60.5; DB 1; Length 86;
Best Local Similarity 29.4%; Pred. No. 27;
Matches 25; Conservative 1; Mismatches 40; Indels 19; Gaps 3;

```

```

RN
  [1]
RN MEDLINE=69008081; PubMed=3170481;
RA Kendall K.J., Cohen S.N.;
RT "Complete nucleotide sequence of the Streptomyces lividans plasmid
RT pJ101 and correlation of the sequence with genetic properties.";
RL J. Bacteriol. 170:4634-4651(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M21778; AAA88413.1; -;
DR PIR; A30924; A30924.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 85 AA; 9339 MW; ABOA558B799A124D CRC64;

Query Match 11.9%; Score 56.5; DB 1; Length 85;
Best Local Similarity 31.0%; Pred. No. 61;
Matches 27; Conservative 9; Mismatches 34; Indels 17; Gaps 6;

```

```

OY 1 PPAATRPGRSRRARNSSSNHLGVAGAPPADPSTLYRD-LPAEDSRGSGDAPTEDDY 60
DB 5 PGPGPGPPPSA-----GEPGPGGRPFPGTPGMQGGGGLGEGGERLPP 53
OY 53 DAPTEDDYWGTVGGEDQGEQMGPG 77
DB 54 PGPGESRTGPPGSGTGRPGPG 78

RESULT 3
Y9KD_STRLI STANDARD; PRT; 85 AA.
AC P22399;
DR 01-AUG-1991 (Rel. 19, Created)
DR 01-AUG-1991 (Rel. 19, Last sequence update)
DR 01-AUG-1991 (Rel. 19, Last annotation update)
DE Hypothetical 9.4 kDa protein (ORF 85).
OS Streptomyces lividans.
OS Plasmid pJ101.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1916;

```

```

OY 7 RPSRA-----RGNSSNHLGVAGAPPADPSTLYRD-LPAEDSRGSGDAPTEDDY 60
DB 4 RPSSSAVALRLRRGRSSGCPV---VGKMPAPADLDRDGYAGRGASVEGEGGRHADR 59
OY 61 WGGYGE---DQR--GEGMCPGAAAC 81
DB 60 HGGPQAQELPHDQRPADDEQC-GQHCR 85

RESULT 4
VESB_BPV4 STANDARD; PRT; 95 AA.
AC P08349;
DR 01-AUG-1988 (Rel. 08, Created)
DR 01-AUG-1988 (Rel. 08, Last sequence update)
DR 01-AUG-1988 (Rel. 08, Last annotation update)
DE Probable 85B protein.
OS Bovine papillomavirus type 4.
OS Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10562;
RN
  [1]
RN MEDLINE=87282264; PubMed=3039043;
RA Patel K.R., Smith K.T., Campo M.S.;
RT "The nucleotide sequence and genome organization of bovine
RT papillomavirus type 4.";
RL J. Gen. Virol. 68:2117-2128(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X05817; -; NOT_ANNOTATED_CDS.
DR PIR; I27129; WSWL5B.
KW Early protein.
SQ SEQUENCE 95 AA; 11026 MW; 84978E1A86576A9F CRC64;

Query Match 11.7%; Score 55.5; DB 1; Length 95;
Best Local Similarity 30.2%; Pred. No. 84;
Matches 19; Conservative 6; Mismatches 31; Indels 7; Gaps 3;

```

DB 4 PHLPETPGAGSGRSLRDRDGHHDRLRGRTPYDETGRYR-VPG-DPREDEGAPPN 61

OY 57 EDD 59
DB 62 GND 64

RESULT 5
ES64_MYCTU STANDARD: PRT: 94 AA.
ID ES64_MYCTU
AC 005300:
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative ESAT-6 like protein 4.
GN RV1198 OR MT1236 OR MTC1364.10.
OS Mycobacterium tuberculosis.
OC Bacteria: Actinobacteria: Actinobacteriia (class): Actinobacteridae:
OC Actinomycetaceae: Corynebacteriaceae: Mycobacteriaceae: Mycobacterium.
NCBI_TaxID=1773;
[1]
SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badoec K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham D., Garsdale J., Hamlin N., Holtgrave S.,
RA Hornsby T., Jagels K., Krogh A., McLean S., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
[2]
SEQUENCE FROM N.A.
RP STRAIN-CDC 1551 / Oshkosh;
RC Fletschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Feldman J., Khouri H., Gill J., Mikula A.,
RA Blahai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z93777; CAB07821.1; -
DR EMBL: AE007000; AAK5493.1; -
DR TIGR: MT1236; -
DR Tuberculin; RV1198; -
KW Hypothetical protein; Complete proteome.
FT CONFLICT 32 32 TAS -> AAG (IN REF. 2).
FT CONFLICT 37 39 TAS -> AAG (IN REF. 2).
SO SEQUENCE 94 AA: 9928 MW: 55858308948846C CRC64;

OY 25 AEAAPPDPSTLYRLDPAEDSRGQGDAPTEDDYMGVGGEDRGQMGCAQO 81
DB 19 AONGLEAEHQAIIRDV-----LTASDFMGAG-----SAACO 51

RESULT 6
OSTR_MOUSE
ID OSTR_MOUSE STANDARD: PRT: 95 AA.
AC P54615;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin-related protein precursor (OC-X) (Nephrocalcin).
GN BGLAP-RS1.
OS Mus musculus (Mouse).
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94117426; PubMed=8288580;
RA Desbois C., Hogue D.A., Karsenty G.;
RT "The mouse osteocalcin gene cluster contains three genes with two
RT separate spatial and temporal patterns of expression."
RL J. Biol. Chem. 269:1183-1190(1994).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=94062692; PubMed=8243336;
RX Rahman S., Oberdorf A., Montecino M., Tanhauser S.M., Lian J.B.,
RA Stein G.S., Lalpis P.J., Stein J.L.;
RT "Multiple copies of the bone-specific osteocalcin gene in mouse and
RT rat."
RL Endocrinology 133:3050-3053(1993).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-129/SVJ; TISSUE=Liver;
RC Yotov W.V., St Arnaud R.;
RA Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY BUT NOT IN BONE.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC BINDING OF CALCIUM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24430; AAA39855.1; -
DR EMBL: S67456; AAB29146.1; -
DR EMBL: U11541; AAB60445.1; -
DR MGP; MGI:88155; Bglap-rs1.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; Vltk_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SMO0069; GLA; 1.
DR PROSITE: PS00011; GU-CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone; Signal;
KW Multigene family.
FT SIGNAL 1 23 PROBABLE.
FT PROPEP 24 49 OSTEOCALCIN-RELATED PROTEIN.
FT CHAIN 50 95 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 62 62 (BY SIMILARITY).
FT MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 69 69 (BY SIMILARITY).
FT MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID
FT DISULFID 68 74 (BY SIMILARITY).
FT BY SIMILARITY.

SO SEQUENCE 95 AA: 10459 MW: 231968AAB55848E CRC64:

Query Match 11.2%; Score 53; DB 1; Length 95;

Best Local Similarity 33.3%; Pred. No. 1.4e+02;

Matches 20; Conservative 2; Mismatches 28; Indels 10; Gaps 2;

QY 3 AAFRPPG-----SARBCNSSNHLYGVAEGAPADPSTLYRDL---PAEDSRGROGG 52
 DB 23 APTGPESDKAFMKQEGNKVYNRLRYLGASVSPDLEPRLEICELDPACDELISNOYG 82

RESULT 7

ES61_MYCTU STANDARD; PRT: 94 AA.

AC P96364; 008120; 008122; Q9L781;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Antigen Meb9.9B (ESAT-6 like protein 1).

(RV1037C OR MT1066 OR MTCY1062.12) AND

(RV3619C OR MT3721 OR MTCY15C10.33 OR MTCY07H7B.03).

OC Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacterium.

OC NCBI_TaxID=1773;

SEQUENCE FROM N.A.

STRAIN=Erismann;

ALDERSON M.R., Bement T., Day C.H., Zhu L., Molesh D., Skeiky Y.A.W.,

Collier R., Lewinson D.M., Reed S.G., Dillon D.C.;

"Expression cloning of an immunodominant family of Mycobacterium

tuberculosis antigens using human CD4(+) T cells.";

J. Exp. Med. 191:551-560(2000).

[2]

SEQUENCE FROM N.A.

STRAIN=H37RV;

MEDLINE=98295987; PubMed=9634230;

Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Davies R., Deillan K., Feltham T., Gentles S., Hamlin N., Holroyd S.,

Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

Suston J.E., Taylor K., Whitehead S., Barrell B.G.;

"Deciphering the biology of Mycobacterium tuberculosis from the

complete genome sequence.";

Nature 393:537-544(1998).

[3]

SEQUENCE FROM N.A.

STRAIN=CDC 1551 / Oshkosh;

Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,

Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,

Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

Bisai W.;

"Whole genome comparison of Mycobacterium tuberculosis clinical and

laboratory strains.";

Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

-1- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL: AF226277; AAF32406.1;

DR EMBL: 292539; CAB06842.1;

DR EMBL: 295436; CAB08822.1;

DR

DR EMBL: AE006989; AAK45317.1;

DR EMBL: AE007171; AAK48082.1;

DR TIGR: MT1066;

DR TIGR: MT3721;

DR TubercuList; RV1037c;

DR TubercuList; RV3619c;

KW Hypothetical protein; Complete proteome.

FT CONFLICT 20 20 Q -> L (IN REF. 1 AND 3; AAK48082).

FT CONFLICT 23 23 S -> L (IN REF. 1 AND 3; AAK48082).

SO SEQUENCE 94 AA: 9833 MW: 8F971BB03789C57E CRC64;

Query Match 10.7%; Score 51; DB 1; Length 94;

Best Local Similarity 26.3%; Pred. No. 2.1e+02;

Matches 15; Conservative 5; Mismatches 13; Indels 24; Gaps 2;

QY 25 AFAAGPAPADPSTLYRDLPAEDSRGROGGDAPEDDYMGSGGEGDGRGQMGACQ 81

DB 19 AAGSLAEHQAIT-----SDVLTASDFWGGAG-----SAAQC 51

RESULT 8

UV15_SCHPO STANDARD; PRT: 87 AA.

AC P40388; 059754;

DT 01-FEB-1995 (Rel. 31, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE UV-induced protein uv15.

UV15 OR SPEC649.04.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OC NCBI_TaxID=4896;

[1]

SEQUENCE FROM N.A.

STRAIN=U741;

Lee J.K., Kim M., Hong S.H., Park S.D.;

Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.

[2]

SEQUENCE FROM N.A.

STRAIN=972;

MEDLINE=21848401; PubMed=11859360;

Wood V., Galliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,

Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodson G.,

Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,

Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

Woodward J., Volkert G., Aert R., Robben J., Grymoprez B.,

Welljens I., Vanthreels E., Rieger M., Schaefer M., Mueller-Auer S.,

Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hubert H.,

Borzym K., Langer I., Beck A., Lehman R., Reinhardt R., Pohl T.M.,

Eger P., Zimmermann W., Wedler H., Mambert R., Purrelle B.,

Goiteau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,

Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,

Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

Domínguez A., Reyvelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

Shpukovskiy G.V., Ussery D., Barrell B.G., Nurse P.,

"The genome sequence of Schizosaccharomyces pombe.";

Nature 415:871-880(2002).

[3]

CHARACTERIZATION.

STRAIN=U741;

MEDLINE=94324936; PubMed=8048925;

RX

RA Lee J.K., Park E.J., Chung H.K., Hong S.H., Joe C.O., Park S.D.;
RT "Isolation of UV-inducible transcripts from Schistosaccharomyces
pombe."
RL Blochem. Biophys. Res. Commun. 202:1113-1119(1994).
CC -1- FUNCTION: REQUIRED FOR THE MAINTENANCE OF VIABILITY OF CELLS IN
STATIONARY PHASE AND IN STARVATION CONDITION.
CC -1- INDUCTION: BY UV LIGHT, ALKYLATING AGENTS AND HEAT SHOCK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: 234300; CAB84070.1; -
DR EMBL: AL023587; CAA19046.1; -
DR Heat shock.
KW CONFLICT 76 87
SEQUENCE 87 AA: 9319 MW: 07A1743C95DD95B7 CRC64;
Query Match 10.5%; Score 50; DB 1; Length 87;
Best Local Similarity 27.2%; Pred. No. 2.4e+02;
Matches 22; Conservative 8; Mismatches 31; Indels 20; Gaps 5;
OY 17 SSNHLGVAAAGAPADSTLY---RDLPEDSRGQGDAPTEDDYWGCGE----- 67
DB 2 SMOGRYGC--DKGYAPAPPOAAGCPNYPPOQNYPOQ--GYAPPGYPOGGRPAQPMYQ 58
OY 68 -----DQGEQMC---PGAAC 80
DB 59 QPQASDPGDLCCGLTGLAC 79
Db 59 QPQASDPGDLCCGLTGLAC 79
RESULT 9
ES66_MYCTU STANDARD: PRT: 94 AA.
ID ES66_MYCTU
AC P95242;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative ESAT-6 like protein 6.
GN RV2346C OR MT2411 OR MTCY98.15C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
[1]
SEQUENCE FROM N.A.
RA STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Bacchok K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Felkell T., Gentles S., Hamlin N., Holroyd S.,
Hornsbly T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Olliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutler S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
[2]
SEQUENCE FROM N.A.
RA STRAIN=CDC 1551 / Oshkosh;
RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolova M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
*Whole genome comparison of Mycobacterium tuberculosis clinical and

RT Laboratory strains."
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: 283860; CAB06161.1; -
DR EMBL: AE007081; AAK46704.1; -
DR TIGR: MT2411; -
DR TubercuList; RV2346C; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 94 AA: 9954 MW: C740622D08B7C8CA CRC64;

Query Match 10.3%; Score 49; DB 1; Length 94;
Best Local Similarity 26.3%; Pred. No. 3.1e+02;
Matches 15; Conservative 5; Mismatches 13; Indels 24; Gaps 2;
OY 25 AEAGAPPADPSTLYRDLPEDSRGQGDAPTEDDYWGCGEGDGRGMCPCGACQ 81
DB 19 AQAGLLEHQAIVRDVLA-----ACDFMGAG-----SVACQ 51
OY 19 AQAGLLEHQAIVRDVLA-----ACDFMGAG-----SVACQ 51

RESULT 10
NEUL_ANSAN STANDARD: PRT: 93 AA.
ID NEUL_ANSAN
AC P35519;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neutrophysin 1 (VLDV-neutrophysin).
DE Neutrophysin 1 (VLDV-neutrophysin).
OS Anser anser anser (Western graylag goose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
OX NCBI_TaxID=8844;
[1]
SEQUENCE.
RA TISSUE=Plutary;
RC MEDLINE=91115507; PubMed=2276874;
RA Michel G., Levy B., Chauvet M.-T., Chauvet J., Acher R.;
RT "Complete amino acid sequence of goose VLDV-neutrophysin. Traces of a
putative gene conversion between promesotocin and provasotocin
genes."
RL Int. J. Pept. Protein Res. 36:457-464(1990).
CC -1- FUNCTION: NEUTROPHYSIN 1 SPECIFICALLY BINDS OXYTOCIN.
CC PIR: A60357; A60357.
DR HSSP; P01180; INPO.
DR InterPro; IPR000981; Neutryp-horm.
DR Pfam; PF00184; hormones; 1.
DR ProDom; PD001676; Neutryp-horm; 1.
DR SMART; SM00003; NH; 1.
FT DISULFID 10 54 BY SIMILARITY.
FT DISULFID 13 27 BY SIMILARITY.
FT DISULFID 28 34 BY SIMILARITY.
FT DISULFID 21 44 BY SIMILARITY.
FT DISULFID 61 74 BY SIMILARITY.
FT DISULFID 68 86 BY SIMILARITY.
FT DISULFID 75 80 BY SIMILARITY.
SQ SEQUENCE 93 AA: 9676 MW: 25C954C9C8C0F23 CRC64;

Query Match 10.2%; Score 48.5; DB 1; Length 93;
Best Local Similarity 24.7%; Pred. No. 3.4e+02;
Matches 20; Conservative 3; Mismatches 29; Indels 29; Gaps 2;
OY 9 GSRAGNSSNHLGVAAAGAPADSTLY-----YRDLPEDSRGQGDAPTEDDYWG 62
DB 14 GPRNRGRGFRICCGELGCVLGPETLRGQESFLPTPESGRKRCGGP----- 64

```

OY 63 GYGEDORGEMCPGACAP 83
Db 65 -----GASCAP 71

RESULT 11
PRPS_HUMAN STANDARD; PRT: 96 AA.
ID P04281;
AC 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DE 01-OCT-1994 (Rel. 30, Last annotation update)
DS Basic proline-rich peptide IB-1.
OS Homo sapiens (Human)
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
SQ MEDLINE=6624335; PubMed=3521730;
Kaufman D., Hofmann T., Bennis A., Keller P.;
"Basic proline-rich proteins from human parotid saliva: complete
covalent structures of proteins IB-1 and IB-6."
Biochemistry 25:2387-2392(1986).
-1- DOMAIN: CONTAINS 21-RESIDUE REPEATS, TWO OF WHICH HAVE INTERNAL
7-RESIDUE REPEATS.
-1- MISCELLANEOUS: THERE ARE NINE BASIC PROLINE-RICH PEPTIDES ISOLATED
FROM THE SALIVA; THIS PEPTIDE IS DESIGNATED IB-1.
DR PIR: A03293; PIHOBI.
KW Phosphorylation; Repeat; Parotid gland; Saliva.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 PHOSPHORYLATION.
FT REPEAT 15 35 APPROXIMATE.
FT REPEAT 36 56 APPROXIMATE.
FT REPEAT 57 77 APPROXIMATE.
SQ SEQUENCE 96 AA: 9530 MW: FF3B3E8DE104F8 CRC64;

Query Match 10.1%; Score 48.5; DB 1; Length 96;
Best Local Similarity 22.6%; Pred. No. 3.5e+02;
Matches 19; Conservative 7; Mismatches 21; Indels 37; Gaps 4;

OY 1 PPAATPGSRARNGSSNHLGYVAEAGAPPDPSTLYRDLPAEDSRGQGDAPTEDDY 60
Db 20 PGAGPGGNGKPGQ-----PPSP-----GKPGGPP----- 46
OY 61 WGYGEGDORGEMCPGACAP 84
Db 47 ---QGNGPQGPPPPPKP-QGPP 66

RESULT 12
GRP7_DAUCA STANDARD; PRT: 96 AA.
ID GRP7_DAUCA
AC P37704;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 01-OCT-1994 (Rel. 30, Last annotation update)
DS Glycine-rich protein DC7.1 precursor.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RA Aleith F., Richter G.;
RT "Gene expression during induction of somatic embryogenesis in carrot
cell suspensions."
RL Planta 183:17-24(1990).
-1- FUNCTION: MAY BE CONNECTED WITH THE INITIATION OF EMBRYOGENESIS
OR WITH THE METABOLIC CHANGES PRODUCED BY THE REMOVAL OF AUXINS.
-1- DEVELOPMENTAL STAGE: TRANSIENTLY EXPRESSED DURING EARLY
EMBRYOGENESIS.

```

```

CC -1- INDUCTION: BY THE REMOVAL OF AUXINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X15706; CA33736.1;
CC DR PIR: S35715; S35715.
CC KW Repeat; Signal.
CC FT SIGNAL 1 25 POTENTIAL.
CC FT CHAIN 26 96 GLYCINE-RICH PROTEIN DC7.1.
CC FT DOMAIN 42 67 2 APPROXIMATE REPEATS OF H-H-G(4,6)-H.
CC FT REPEAT 42 50 1.
CC FT REPEAT 61 67 2.
CC SQ SEQUENCE 96 AA: 9319 MW: 7C00D44637B7A364 CRC64;

Query Match 10.1%; Score 48; DB 1; Length 96;
Best Local Similarity 27.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 4; Mismatches 38; Indels 12; Gaps 2;

OY 21 LYGAEAAGAPPDPSTLYRDL-----AEDSRGQGDAPTEDDYWGYGEGDORG----- 71
Db 8 LLGLIAFALLISSEVAARDLSETTGGASLDGHHGGGCGGSGGSGHGGG 67
OY 72 ---EOMCPGACQA 82
Db 68 GCGHHYCHGSCSA 81

RESULT 13
VE4_HP28 STANDARD; PRT: 99 AA.
ID VE4_HP28
AC P51896;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DS Probable E4 protein.
OS Human papillomavirus type 28.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
CC Papillomavirus.
CC OX NCBI_TaxID=37111;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA Delius H.;
CC RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U31783; AAA79426.1; ALT_INIT.
CC DR InterPro: IPR003861; Papilloma_E4.
CC DR Pfam: PF02711; Pap_E4; 1.
CC KW Early protein.
CC SQ SEQUENCE 99 AA: 11112 MW: 4FF84D4D52DBC51E CRC64;

Query Match 10.1%; Score 48; DB 1; Length 99;
Best Local Similarity 32.1%; Pred. No. 4e+02;
Matches 9; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 7 PPGSRARNGSSNHLGYVAEAGAPPDP 34
Db 34 RPKDRSKNDSRHSSTGSSSDSTPKPP 61

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 32.1111 Seconds
(without alignments)
254.473 Million cell updates/sec

Title: us-09-972-546-2_COPY_311_395

Perfect score: 475
Sequence: 1 PPAAPTRPGSRARGNSSNNH.....GEGDQGEQMGCPGAACQAPD 85

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

1 number of hits satisfying chosen parameters: 28758

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60.5	12.7	86	2 S00802	collagen alpha 1(X
2	59.5	12.5	74	2 S68269	Y-box binding prot
3	59.5	12.5	97	2 T35099	hypothetical prote
4	56.5	11.9	60	2 B70809	hypothetical prote
5	56.5	11.9	85	2 A50924	hypothetical 10.8k
6	55.5	11.7	95	1 W5WL5B	E5B protein - bovi
7	55	11.6	94	2 D70608	hypothetical prote
8	53	11.2	89	2 S55780	Hox C6 protein - e
9	53	11.2	95	2 I67413	osteocalcin - mus
10	53	11.2	95	2 I61188	osteocalcin-relate
11	52.5	11.1	96	2 S77116	hypothetical prote
12	51.5	10.8	96	2 B44962	collagen COLA4 - p
13	51	10.7	61	2 B95850	hypothetical prote
14	51	10.7	89	2 T19805	hypothetical prote
15	51	10.7	94	2 D70560	hypothetical prote
16	50	10.5	59	2 T05022	hypothetical prote
17	50	10.5	80	2 S54845	GTP cyclonhydrolase
18	50	10.5	87	2 T40597	hypothetical prote
19	50	10.5	93	2 I45876	collagen alpha 1(I
20	50	10.5	93	2 AB0740	probable excisiona
21	49.5	10.4	72	2 T67566	hypothetical prote
22	49.5	10.4	97	2 B84746	hypothetical prote
23	49.5	10.4	100	2 S20482	hypothetical prote
24	49	10.3	45	2 G82812	hypothetical prote
25	49	10.3	74	2 C23734	insulin-like growt
26	49	10.3	78	2 B83446	hypothetical prote
27	49	10.3	73	2 C70662	hypothetical prote
28	48.5	10.2	73	2 E70833	hypothetical prote
29	48.5	10.2	93	2 A60357	neurophysin I - go

30	48.5	10.2	98	2 S36048	Ig lambda chain -
31	48	10.1	81	2 T35731	hypothetical prote
32	48	10.1	96	2 S35715	glycine-rich prote
33	48	10.1	96	2 B81047	hypothetical prote
34	47.5	10.0	50	2 A29789	mucin - sheep (tra
35	47.5	10.0	92	2 A38947	oncofetal-laminin
36	47.5	10.0	99	2 S44023	RNA-binding protei
37	47	9.9	63	2 PC4356	dynorphin precurs
38	47	9.9	83	2 C82519	hypothetical prote
39	47	9.9	85	2 JH0786	transcription fact
40	47	9.9	90	2 S34247	hypothetical prote
41	47	9.9	96	2 F75414	Glu-tRNA(Glu) amid
42	46.5	9.8	92	2 S41723	hypothetical prote
43	46	9.7	59	2 D82750	hypothetical prote
44	46	9.7	66	2 T35419	small hypothetical
45	46	9.7	71	2 T09239	nlfu protein - Fra

ALIGNMENTS

```

RESULT 1
S00802
collagen alpha 1(XII) chain - bovine (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 08-Dec-1995
C:Accession: S00802
R:Dublet, B.; Dixon, E.; de Miguel, E.; van der Rest, M.
FEBS Lett. 233, 177-180, 1988
A>Title: Bovine type XII collagen: amino acid sequence of a 10 kDa pepsin fragment fr
A:Reference number: S00802; MUID:88255287; PMID:3133242
A:Accession: S00802
A:Molecule type: protein
A:Residues: 1-15;16-48;49-86 <DUB>
C:Superfamily: collagen alpha 1(XII) chain; fibronectin type III repeat homology; von
C:Keywords: cell binding; coiled coil; extracellular matrix; glycoprotein; hydroxypro
F:6,9,12,18,24,27,30,42,51,54,65,74,77,80/Modified site: hydroxyproline (Pro) #status

Query Match          12.7%  Score 60.5; DB 2; Length 86;
Best Local Similarity 29.4%; Pred. No. 53;
Matches 25; Conservative 1; Mismatches 40; Indels 19; Gaps 3;

OY 1 PPAAPTRPGSRARGNSSNNHLYVAAGAPPADPSTL-----YRDLPAE-DSKRGCG 52
DB 5 PPGPGPGGSA-----GEPGPGRPGFPGTGMQPGQGERGLPGEXGERGLRCP 53
OY 53 DAPTEDDYWGCGEGEDQGEQMGCPG 77
DB 54 PGPQGESRTGPPGSGSRGPPGPG 78

RESULT 2
S68269
Y-box binding protein YB-1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68269
R:Makino, Y.; Ohga, T.; Toh, S.; Koike, K.; Okumura, K.; Wada, M.; Kuwano, M.; Kohno,
Nucleic Acids Res. 24, 1873-1878, 1996
A>Title: Structural and functional analysis of the human Y-box binding protein (YB-1)
A:Reference number: S68269; MUID:96226173; PMID:8657568
A:Accession: S68269
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-74 <MAX>
A:Cross-references: EMBL:X96666; NID:91403348; PIDN:CA65446.1; PID:91403349
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, March 1996
C:Genetics:
A:Gene: YB-1
C:Superfamily: Y box-binding protein 1; cold shock domain homology

Query Match          12.5%  Score 59.5; DB 2; Length 74;
Best Local Similarity 32.1%; Pred. No. 56;

```

Matches 26; Conservative 7; Mismatches 29; Indels 19; Gaps 4;

Oy 1 PPAAPTRPGSGARAGNSSNNHLYGVAEACAPPADPSTLYRDLPAEDSRGRGGDAFTEDDY 60
||||| : : : : :
Db 10 PPAAV--PPAAPL--SAADRKPTGTGSGAGSGCGPLTNSAAPA-----GDKKVI--- 55

Oy 61 WGYGGEORGORGFOMCPGAACQ 81
|||:| | |||
Db 56 -----GEDRGGRGMCPRAAQ 71

RESULT 3

T35099
hypothetical protein SC4G6.39c - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: J35099
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, July 1999
Reference number: Z21567
Accession: J35099
Annotation: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-97 <SE>
A:Cross-references: EMBL:AL096884; PIDN:CAB51462.1; GSPDB:GN00070; SCOEDB:SC4G6.39c
C:Genetics:
A:Gene: SCOEDB:SC4G6.39c

Query Match 12.5%; Score 59.5; DB 2; Length 97;
Best Local Similarity 36.7%; Pred. No. 73;
Matches 18; Conservative 4; Mismatches 24; Indels 3; Gaps 2;

Oy 31 PADPSTLYRDLPADPSRGRGGDAFTEDDYWGYGGEORGORGFOMCPGAA 79
||||| : : : : :
Db 8 PADPGAVLRHPHGGLDVRSSQ--QARAHDIVDG-AGTAQQPQRCLGVS 53

RESULT 4

B70809
hypothetical protein Rv0810c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70809
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, M.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Submitted to the EMBL Data Library, Jul 1998
Reference number: Z37544, 1998
Accession: U08393, 537-544, 1998
Annotation: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
Accession number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70809
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-60 <COL>
A:Cross-references: GB:AL022004; GB:AL123456; NID:q3261550; PIDN:CAA17616.1; PID:e125395
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0810c

Query Match 11.9%; Score 56.5; DB 2; Length 60;
Best Local Similarity 27.5%; Pred. No. 86;
Matches 14; Conservative 8; Mismatches 28; Indels 1; Gaps 1;

Oy 11 RAKGNSSNNHLYGVAEACAPPADPSTLYRDLPAEDSRGRGGDAFTEDDY 61
|||:| | |||
Db 3 RORAKAKOTKVARELKYSPOTFORLORELSTGT-DRLDDGPSDDSW 52

RESULT 5

A30924
hypothetical 10.8k protein (ORF 85) - Streptomyces lividans plasmid pIJ101
C:Species: Streptomyces lividans

```

CjDate: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 18-Jun-1993
CjAccession: A30924
R:Kendall, K.J.; Cohen, S.N.
J. Bacteriol. 170, 4634-4631, 1988
A>Title: Complete nucleotide sequence of the Streptomyces lividans plasmid pU101 and
A:Reference number: A91888; MUID:8908081; PMID:3170481
A:Accession: A30924
A:Molecule type: DNA
C:Genetics:
A:Residues: 1-85 <KEN>
C:genome: plasmid

Query Match          11.9%; Score 56.5; DB 2; Length 85;
Best Local Similarity 31.0%; Pred. No. 1.2e+02;
Matches 27; Conservative 9; Mismatches 34; Indels 17; Gaps 6;

QY      7  RRGSA-----RGNSSNHLYGVAEGAPPADPSTLYRD-LPAEDSRGGGDAPTEDDY 60
         ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      4  RSSSAVRLRRGRSSGPPIV---VGMPAPADLDLRDGYAGRGRASVEGEGRHADR 59
         ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      61  WGVGGE---DQR--GEOMCPGAACQ 81
         ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      60  HCGQAQLPHDQRPADQEQ-CQHCR 85
         ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 6
MSL5B
E5B protein - bovine papillomavirus type 4
CjSpecies: bovine papillomavirus type 4
CjDate: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 24-Feb-1994
CjAccession: I27129
R:Patel, K.R.; Smith, K.T.; Campo, M.S.
J. Gen. Virol. 68, 2117-2128, 1987
A>Title: The nucleotide sequence and genome organization of bovine papillomavirus typ
A:Reference number: A92795; MUID:87282264; PMID:3039043
A:Accession: I27129
A:Molecule type: DNA
A:Residues: 1-95 <PAT>
CjSuperfamily: bovine papillomavirus type 4 E5B protein
CjKeywords: early protein

Query Match          11.7%; Score 55.5; DB 1; Length 95;
Best Local Similarity 30.2%; Pred. No. 1.7e+02;
Matches 19; Conservative 6; Mismatches 31; Indels 7; Gaps 3;

QY      2  PAAPRPGSRARGNS-----SSNHLGYAEAGAPPADPSTLYRDLPAEDSGRGGDAPT 56
         ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      4  PHLPTPGAAGSKRSRLDRDHGHDLRLRKRTPVDETGRY-VPG-DPREDEGAPPN 61
         ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      57  EDD 59
         ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      62  GND 64

RESULT 7
D70608
hypothetical protein RV1198 - Mycobacterium tuberculosis (strain H37RV)
CjSpecies: Mycobacterium tuberculosis
CjDate: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
CjAccession: D70608
R:Cole, S.T.; Brooch, R.; Parkhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon
J.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
R.; Rajandrem, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70608
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-94 <COL>
A:Cross-references: GB:293777; GB:AL123456; NID:g3261726; PIDN:CAB07821.1; PID:e1107
A:Experimental source: strain H37RV

```

C:Genetics:
A:Gene: Rv1198

Query Match
Best Local Similarity 28.1%; Score 55; DB 2; Length 94;
Matches 16; Conservative 5; Mismatches 12; Indels 24; Gaps 2;

OY 25 AEGAPADPSTLYRDLPAEDSRGOGDAPTEDDYWGCGEDRGEGMCGAACQ 81
DB 19 AAGGLEAEHQATIRDV-----LNASDFMGACG-----SAACQ 51

RESULT 8

Hox C6 protein - eastern newt (fragment)
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 07-May-1999
C:Accession: S55780
R:Savard, P.; Tremblay, M.
J. Mol. Biol. 249, 879-889, 1995
Title: Differential regulation of Hox C6 in the appendages of adult urodeles and anura
A:Accession: S55780; MUID:95311312; PMID:7791214
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-89 <SAV>

Query Match
Best Local Similarity 35.3%; Score 53; DB 2; Length 89;
Matches 11; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

OY 44 EDSRGOGDAPTEDDYWGCGEDRGEGM 74
DB 53 EKRNKRGCDREDEALGVGYCTGRRGRQI 83

RESULT 9

Osteocalcin - mouse
C:Species: Mus sp. (mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: I67413
R:Rahman, S.; Oberdorfer, A.; Montecino, M.; Tanhauser, S.M.; Llan, J.B.; Stein, G.S.; Lel
Endocrinology 139, 3050-3053, 1993
A:Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.
A:Reference number: I55275; MUID:94062692; PMID:8243336
A:Accession: I67413
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-95 <RES>

A:Cross-references: GB:S67456; NID:9456855; PIDN:AA829146.1; PID:9456857
A:Introns: 22/1; 33/1; 52/2; 72/2
C:Superfamily: osteocalcin

Query Match
Best Local Similarity 33.3%; Score 53; DB 2; Length 95;
Matches 20; Conservative 2; Mismatches 28; Indels 10; Gaps 2;

OY 3 AAPTRG-----SRARSSSNHLYGVAAGAPADPSTLYRDL---PAEDSRGOGG 52
DB 23 APTGPESDKAFMSKQEGNKVNRRLRYLGASVSPDLPTRELCELDPACDELNSQYG 82

RESULT 10

Osteocalcin-related protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C:Accession: I61188; I49073
R:Desbois, C.; Hogue, D.A.; Karsenty, G.
J. Biol. Chem. 269, 1183-1190, 1994
A:Title: The mouse osteocalcin gene cluster contains three genes with two separate spatial

A:Reference number: A49871; MUID:94117426; PMID:8288580
A:Accession: I61188

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:I24430; NID:9455454; PIDN:AAA39855.1; PID:9455455

A:Accession: I49073

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: EMBL:U11541; NID:9508297; PIDN:AA860445.1; PID:9508298

C:Genetics:

A:Introns: 22/1; 33/1; 52/2

C:Superfamily: osteocalcin

Query Match
Best Local Similarity 33.3%; Score 53; DB 2; Length 95;
Matches 20; Conservative 2; Mismatches 28; Indels 10; Gaps 2;

OY 3 AAPTRG-----SRARSSSNHLYGVAAGAPADPSTLYRDL---PAEDSRGOGG 52
DB 23 APTGPESDKAFMSKQEGNKVNRRLRYLGASVSPDLPTRELCELDPACDELNSQYG 82

RESULT 11

hypothetical protein ssl3383 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77116
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-90 <KAN>

A:Cross-references: EMBL:D90908; GB:AB001339; NID:91652725; PIDN:BA17674.1; PID:d101
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match
Best Local Similarity 32.9%; Score 52.5; DB 2; Length 90;
Matches 23; Conservative 7; Mismatches 19; Indels 21; Gaps 5;

OY 22 YGVAEGAPP---ADPST-----LYRDLPAEDSRGOGG---DAPTDDYWGCGG 66
DB 27 FGMALRSAPPRAVADPSVPTPEQSGGLFR---SQGSFPQAGNNLSQGRDE---NHGG 80

RESULT 12

B44982
collagen COL4A - pig roundworm
C:Species: Ascaris suum (pig roundworm)
C:Date: 14-May-1993 #sequence_revision 30-Sep-1993 #text_change 20-Sep-1999
C:Accession: B44982
R:Kingston, I.B.; Wainwright, S.M.; Cooper, D.
Mol. Biochem. Parasitol. 37, 137-146, 1989
A:Title: Comparison of collagen gene sequences in Ascaris suum and Caenorhabditis ele
A:Reference number: A44982; MUID:90136710; PMID:2482444
A:Accession: B44982
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-96 <KIN>
A:Cross-references: GB:J04655; NID:9159654; PIDN:AAA29370.1; PID:9159655
C:Superfamily: unassigned collagens

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 73.037 Seconds
(without alignments)
155.076 Million cell updates/sec

Title: US-09-972-546-2_COPY_311_395
Perfect score: 475
Sequence: 1 PPAPTRPGSRARCNSNNH.....GEDORGEQMPGACQAPPD 85

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
1 number of hits satisfying chosen parameters: 180751

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.5	15.5	88	22	ABG08602
2	69.5	14.6	80	22	AAO10768
3	68.5	14.4	77	22	AAU22202
4	67	14.1	62	22	ABR17420
5	65.5	13.8	77	22	AAU62866
6	65	13.7	56	23	ABP35282
7	65	13.7	94	22	ABR57388
8	65	13.7	94	23	ABG55366
9	64.5	13.6	96	21	AAZ6174
10	64	13.5	68	22	AAU22312

11	62	13.1	85	22	AAM84780
12	61	12.8	51	22	AAU57626
13	61	12.8	91	22	AAU50651
14	59.5	12.5	99	22	AAO08818
15	59	12.4	68	22	AAU59903
16	59	12.4	88	22	AAM06671
17	58.5	12.3	91	21	AAV84550
18	57.5	12.1	89	23	ABP04725
19	57	12.0	92	22	AAU00828
20	57	12.0	95	21	AAAG4126
21	56.5	11.9	61	21	AAAB03850
22	56.5	11.9	67	23	ABP33674
23	56.5	11.9	74	22	AAU43366
24	56.5	11.9	80	22	AAO11030
25	56	11.8	50	22	AAU48331
26	56	11.8	77	21	AAV56033
27	56	11.8	84	21	AAAG40969
28	56	11.8	97	21	AAAG40991
29	55.5	11.7	63	22	AAU43588
30	55.5	11.7	66	21	AAAG40982
31	55.5	11.7	76	22	AAO00304
32	55.5	11.7	86	22	AAAB93053
33	55	11.6	64	21	AAAG16078
34	55	11.6	64	21	AAAG51128
35	55	11.6	65	21	AAAG16077
36	55	11.6	68	22	ABG11143
37	55	11.6	75	22	AAU40283
38	55	11.6	81	23	ABP10567
39	55	11.6	86	18	AAAB15789
40	55	11.6	88	22	ABG01422
41	55	11.6	90	16	AAAR2476
42	55	11.6	90	22	ABR38623
43	55	11.6	90	22	ABR33718
44	55	11.6	90	22	AAAM59252
45	55	11.6	90	22	AAAM71792

ALIGNMENTS

RESULT 1
ABG08602
ID ABG08602 standard; Protein: 88 AA.
AC ABG08602;
XX
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE Novel human diagnostic protein #8593.
XX
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200175067-A2.
XX
XX
PD 11-OCT-2001.
XX
XX
PP 30-MAR-2001; 2001WO-US08631.
XX
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX
DR WPI: 2001-639362/73.
XX
XX
PT N-PSDB: AAS72789.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID NO 38961; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SO Sequence 88 AA:
 Query Match 15.5%; Score 73.5; DB 22; Length 88;
 Best Local Similarity 32.9%; Pred. No. 1.8;
 Matches 25; Conservative 6; Mismatches 18; Indels 27; Gaps 4;
 OY 2 PAAPTR-PCSRARGN-----SSNNHLYGVADACAP-----PADPST 36
 Db 7 PAAPTRGPGARVRGHPPCAPPPIRKAAATAPKLGCGVGAPSPKQTMREKALHVAAPATPT 66
 OY 37 LYRDLPAEDSRGRCG 52
 Db 67 GYK--PPGGGRTQEG 80
 RESULT 2
 AA010768
 ID AA010768 standard; Protein; 80 AA.
 XX
 AC AA010768;
 DE 06-NOV-2001 (first entry)
 XX Human polypeptide SEQ ID NO 24660.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.

DR N-PSDB: AA190699.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 PS Claim 20; SEQ ID NO 24660; 1399pp + Sequence listing; English.
 XX
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SO Sequence 80 AA:
 Query Match 14.6%; Score 69.5; DB 22; Length 80;
 Best Local Similarity 32.4%; Pred. No. 4.3;
 Matches 22; Conservative 3; Mismatches 24; Indels 19; Gaps 3;
 OY 21 LYGVADACAPPADPSTLYRDLPAEDSRGRCGDART-----EDYVGGGGEQ 69
 Db 7 LIGV-QPGAPPDST-----ASGSGARGGVPVTLGGNTGSRKWDPCWSONSPSS 58
 OY 70 RGEOMCPG 77
 Db 59 DSGKPPPG 66
 RESULT 3
 AA022202
 ID AA022202 standard; Protein; 77 AA.
 XX
 AC AA022202;
 DE 18-DEC-2001 (first entry)
 XX Human cardiovascular system antigen polypeptide SEQ ID NO 976.
 KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
 KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
 KW antirheumatic; antiproliferative; cytotatic; cardiant; neuroprotective;
 KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
 KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasia;
 KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
 KW cerebrovascular disorder; nervous system disorder; bacterial infection;
 KW fungal infection; viral infection; ocular disorder; endocrine disorder;
 KW gastrointestinal disorder; renal disorder; respiratory disorder;
 KW wound healing; skin aging; organ transplant; tissue regeneration;
 KW anti-infertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200155321-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01340.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0233081.
 PR 12-SEP-2000; 2000US-0233968.
 PR 14-SEP-2000; 2000US-0233969.
 PR 14-SEP-2000; 2000US-023398.
 PR 14-SEP-2000; 2000US-023399.
 PR 14-SEP-2000; 2000US-0234001.
 PR 14-SEP-2000; 2000US-0234001.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251980.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451930/48.

N-PSDB; AAS35476.

New cardiovascular system related polynucleotides and polypeptides,
 useful for diagnosing, treating and/or preventing disorders of the
 cardiovascular system -

Claim 11; SEQ ID No 976; 674bp; English.

Sequences AU21852-AU22466 represent the cardiovascular system antigen

PR 20-OCT-2000: 2000US-0241826.
 PR 20-OCT-2000: 2000US-0242221.
 PR 01-NOV-2000: 2000US-0244617.
 PR 08-NOV-2000: 2000US-0246474.
 PR 08-NOV-2000: 2000US-0246475.
 PR 08-NOV-2000: 2000US-0246476.
 PR 08-NOV-2000: 2000US-0246477.
 PR 08-NOV-2000: 2000US-0246478.
 PR 08-NOV-2000: 2000US-0246523.
 PR 08-NOV-2000: 2000US-0246524.
 PR 08-NOV-2000: 2000US-0246525.
 PR 08-NOV-2000: 2000US-0246526.
 PR 08-NOV-2000: 2000US-0246527.
 PR 08-NOV-2000: 2000US-0246528.
 PR 08-NOV-2000: 2000US-0246532.
 PR 08-NOV-2000: 2000US-0246609.
 PR 08-NOV-2000: 2000US-0246610.
 PR 08-NOV-2000: 2000US-0246611.
 PR 17-NOV-2000: 2000US-0246613.
 PR 17-NOV-2000: 2000US-0249207.
 PR 17-NOV-2000: 2000US-0249208.
 PR 17-NOV-2000: 2000US-0249209.
 PR 17-NOV-2000: 2000US-0249210.
 PR 17-NOV-2000: 2000US-0249211.
 PR 17-NOV-2000: 2000US-0249212.
 PR 17-NOV-2000: 2000US-0249213.
 PR 17-NOV-2000: 2000US-0249214.
 PR 17-NOV-2000: 2000US-0249215.
 PR 17-NOV-2000: 2000US-0249216.
 PR 17-NOV-2000: 2000US-0249217.
 PR 17-NOV-2000: 2000US-0249218.
 PR 17-NOV-2000: 2000US-0249244.
 PR 17-NOV-2000: 2000US-0249245.
 PR 17-NOV-2000: 2000US-0249264.
 PR 17-NOV-2000: 2000US-0249265.
 PR 17-NOV-2000: 2000US-0249297.
 PR 17-NOV-2000: 2000US-0249299.
 PR 01-DEC-2000: 2000US-0249300.
 PR 01-DEC-2000: 2000US-0250391.
 PR 05-DEC-2000: 2000US-0251160.
 PR 05-DEC-2000: 2000US-0251030.
 PR 05-DEC-2000: 2000US-0251988.
 PR 06-DEC-2000: 2000US-0256719.
 PR 06-DEC-2000: 2000US-0251479.
 PR 08-DEC-2000: 2000US-0251856.
 PR 08-DEC-2000: 2000US-0251868.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251989.
 PR 11-DEC-2000: 2000US-0254097.
 PR 05-JAN-2001: 2001US-0255678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-541565/60.
 DR N-PSDB: ABA13746.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 PS
 XX Claim 11: SEQ ID NO 6077; 1701bp + Sequence Listing: English.
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 62 AA:
 Query Match 14.1%; Score 67; DB 22; Length 62;
 Best Local Similarity 32.3%; Pred. No. 6;
 Matches 20; Conservative 10; Mismatches 20; Indels 12; Gaps 3;
 QY 18 SNHLGVAVAGAPADPSTLYRDLPAEDSRGSGDAPT---EDDYMGYGGEIDRGDM 74
 DB 4 SSHLMGLAKKGRKMG---LDMPTPAKSPDPREGGCSFLGAGCKGA-----ORGEAL 54
 QY 75 CP 76
 DB 55 SP 56
 RESULT 5
 ID AUA62866 standard; Protein: 77 AA.
 AC AUA62866;
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #23762.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN MO200181581-A2.
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001: 2001WO-US12865.
 XX
 PR 21-APR-2000: 2000US-199047P.
 PR 02-JUN-2000: 2000US-208841P.
 PR 07-JUL-2000: 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L.maisonneuve J, Zhang Y, Jen S, Carter D;
 DR WPI: 2001-616774/71.
 DR N-PSDB: AASS9630.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 PS
 XX Example 1: SEQ ID NO 24061; 1069pp; English.
 CC Sequences AUA3105-AUA68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of *P. acnes* in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for *P. acnes* proteins. These antibodies can be used to
 CC downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pct_sequences.

CC Sequence 77 AA:

Query Match 13.8%; Score 65.5; DB 22; Length 77;

Best Local Similarity 28.3%; Pred. No. 11; Mismatches 26; Conservative 4; Indels 21; Gaps 5;

3 AAPTTP-GSRAKSSNHLTVGVAAGAPPDPTLYTDLAEBSRGSGDAPTEDY 61
 6 ATRNPASSCKCDWMDKMLGTVRES-----YRD-----W 35

QY 62 GYGGEDRGEGM---CPGAA-----COAP 83
 DB 36 SSGHGEPSPGDLPISTCRMAATSLRTFCQH 67

RESULT 6

ABP35282 ID ABP35282 standard; Protein: 56 AA.

XX ABP35282;

DT 08-JUL-2002 (first entry)

DE Human ORF425 protein, SEQ ID NO:8510.

XX Human; ORF: open reading frame; ORFX: drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnary;
 KW vasotrophic; antipsoriatic; antidiabetic; cytosolic; noctropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antihypoid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX Homo sapiens.

OS WO200190366-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001MO-US17076.

XX 24-MAY-2000; 2000US-206690P.

XX (CURA-) CURAGEN CORP.

XX Leach MD, Shinkets RA;

XX WPI; 2002-106200/14.

XX N-PSDB; ABN79308.

PT Novel human polypeptides and polynucleotides useful for diagnosing,
 preventing and treating cardiovascular disease, neurodegenerative,

PT hyperproliferative disorders and disorders related to organ
 PT transplantation

PS Claim 10: page 2370; 2508bp; English.

XX Sequences ABP35282-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX polynucleotides and
 CC specific for ORFX proteins, methods for detecting ORFX polynucleotides and
 CC polypeptides, methods of screening individuals for a predisposition to an
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokine activity, haemostatic activity, tumour inhibition activity,
 CC receptor/ligand, antiinflammatory activity, chemotactic activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as epilepsy and Alzheimer's disease,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX Sequence 56 AA:

Query Match 13.7%; Score 65; DB 23; Length 56;

Best Local Similarity 39.2%; Pred. No. 8.9; Mismatches 20; Conservative 3; Indels 8; Gaps 3;

QY 40 DLPAPDSRGRG---GDAPTEDYWGTVG---GEDORGEQMPGACQAP 83

DB 3 DPLPSGSGQWPSLGDGPSA-LIMGFLQVSGEQLRSYLDPRGASRLAP 52

RESULT 7

ABAB7388 ID AABAB7388 standard; Protein: 94 AA.

XX AABAB7388;

DT 22-MAY-2001 (first entry)

DE Human gene 47 encoded secreted protein HFXDK20, SEQ ID NO:129.

XX Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angioinfectious disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnary;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.

OS Homo sapiens.
XX
PN WO200118022-A1.
XX
PD 15-MAR-2001.
XX
PP 31-AUG-2000; 2000WO-US24008.
XX
PR 03-SEP-1999; 99US-0152315.
PR 03-SEP-1999; 99US-0152317.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI NI J, Baker KP, Blise CE, Piscella M, Komatsoulis GA, Rosen CA,
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW,
PI Moore PA, Shi Y, Wei Y, Florence KA,
DR WPI: 2001-203081/20.
DR N-PSDB: AAF91904.

Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -
 PT
 XX
 PS
 Claim 11: Page 559: 607pp: English.
 XY

AAFG91858-AAFG91929 represent cDNAs corresponding to 52 human secreted proteins genes, and AAB87342-AAB87413 represent the proteins they encode. AAB87414-AAB87454 represent human secreted protein fragments. The genes and their corresponding secreted proteins are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 52 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein of the invention.

SQ Sequence 94 AA;

Query Match	13.7%	Score 65;	DB 22;	Length 94;
Best Local Similarity	31.9%	Pred. NO. 16;		
Matches 23; Conservative	6;	Mismatches 15;	Indels 28;	Gaps 4;

QY	24	VAEAGAPPP-----DSSTLYRDLPAEDSRGGCAPTEDD-----YMGVGYGEQGR	72
		: : : : : : : : : : : : : : : : :	
Db	37	VSTSPPEPATRCRDPDESELYRD-----PGEAPLENDQAEKGAAHDEGCHPGRD-----	82
QY	73	QMGCGAACGAPP	84
		: : : :	
Db	83	---PMGARRGCP	91

RESULT 8
ABG65366

ID	ABG65366 standard; Protein: 94 AA.
XX	
AC	ABG65366;
XX	
XX	
DT	27-AUG-2002 (first entry)
XX	
XX	
DE	Human albumin fusion protein #2041.
XX	

KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytotoxic; antifertility; antinflammatory; anticancer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.

05	Homo sapiens.
05	Synthetic.

PN WO200177137-A1
XY

PD 18-OCT-2001

PF 12-APR-2001; 2001WO-US11988.
YY

PR 12-APR-2000; 2000US-229358P.
PR 05-APR-2000; 2000US-100000P.

PR 21-DEC-2000; 2000US-256931P.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
YY

PI Rosen CA, Haseltine WA;
 YX

DR WPI; 2002-010886/01
xy

PT New fusion protein for treating disease e.g. diabetes comprises an albumin fused to a therapeutic protein -

PS Claim 1; Page 1949-1950; 2102pp; English.

The present invention relates to albumin fusion proteins comprising a therapeutic protein X and human albumin (HA, also known as human serum albumin, HSA). The proteins are useful for treating a disease or disorder that may be modulated by therapeutic protein X. The albumin extends the shelf-life of protein X, and may increase its biological in vitro/in vivo activity. The protein is useful for treating and diagnosing disorders such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes), haematopoietic disorders, neural disorders (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis, schizophrenia), and connective disorders (e.g. osteoporosis, arthritis). AB663326-AB665518 represent albumin fusion proteins of the invention.

Sequence 94 AA;

Query Match	13.7%	Score 65;	DB 23;	Length 94;
Best Local Similarity	31.9%;	Pred. No. 16;		
Matches 23; Conservative	6;	Mismatches 15;	Indels 28;	Gaps 4;

QY	73 QMCGAQAQAP 84	24 VAEAGAPPA----	37 VSTSPPPATRCRDESELYRD-----	82 Db
QY	83 ---PWGARGRP 91	25 DESTLYRDLPAEDSRGQGAAPED-----	38 PGEAPLEADQAEAGAAHESGHPGRD-----	83 Db
QY		26 YWCGYGGEGORGE 72 QY		

RESULT 9
AAAG26174

ID AAG26174 standard; Protein: 96 AA.
XX AAG26174;
AC
XX 17-OCT-2000 (first entry)
DT
XX
DE Zea mays protein fragment SEQ ID NO: 30531.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence; corn.
XX
XX Zea mays subsp. mays.
XX EPI033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135324.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.68; Score 64.5; DB 21; Length 96;
Best Local Similarity 33.98; Pred. No. 18;

Matches 20; Conservative 6; Mismatches 26; Indels 7; Gaps 3;

OY 1 PPAATRPGRARRGSSNHLGVAEAGAPPADPSTLYRDLPAEDSRGSGDAPTEED 59
DB 8 PPPPTR---RLKSLPASGHL--ASAPGSPPATOTTSVRS--PRDRTGDHGGAHRRPD 59

RESULT 10
AAU22312
ID AAU22312 standard; Protein: 68 AA.
XX
AC AAU22312;

XX 17-DEC-2001 (first entry)
XX Human cardiovascular system antigen polypeptide SRQ ID No 1086.
DE

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW Chicken; sheep; immunosuppressive; antiarthritis; vasotropic; dog;
KW antihemetic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-fertility.
OS Homo sapiens.
XX WO20015321-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01340.
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241825.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI. 2001-451930/48.
 DR N-PSDB; AAS35586.
 DR
 PT New cardiovascular system related polynucleotides and polypeptides,
 PT useful for diagnosing, treating and/or preventing disorders of the
 PT cardiovascular system -
 PR
 PS Claim 11; SEQ ID No 1086; 674pp; English.

CC Sequences AAU21852-AAU22466 represent the cardiovascular system antigen
 CC polypeptides of the invention. Cardiovascular system antigens and their
 CC associated polynucleotides are useful in the diagnosis, treatment and
 CC prevention of various types of disorders in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. A pathological condition
 CC can be determined by detecting the presence or absence of a mutation in a
 CC cardiovascular system antigen polynucleotide. The treatable disorders
 CC include autoimmune diseases such as rheumatoid arthritis,
 CC hyperproliferative disorders such as neoplasms of the breast or liver,
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular
 CC disorders, such as cerebral ischaemia, nervous system disorders such as
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi;
 CC ocular disorders such as corneal infection, endocrine disorders such as
 CC premature labour and infertility, gastrointestinal disorders such as
 CC Crohn's disease, renal disorders such as glomerulonephritis and
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, to regenerate tissues and in
 CC chemotaxis.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

Query Match 13 5%; Score 64; DB 22; Length 68;

Best Local Similarity 35.2%; Pred. No. 14; Mismatches 16; Gaps 3;

Matches 19; Conservative 1; Mismatches 18; Indels 16; Gaps 3;
 Oy 30 PPADPSTLYRDLPAEDSRGSGDAFTEDDYGYGEGDQRGEDQMGPGACGAP 83
 Db 1 PVPSP--HEDLPA-----DKQECTCGEGY-----RWNGSCPAPCGCP 38

RESULT 11

AA084780
 AA084780 standard; Protein; 85 AA.

AC AA084780;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen SEQ ID NO:12373.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

OS cytototoxic; gene therapy; vaccine; metastasis.

XX Homo sapiens.

PN WO200157182-A2.

FE	17-JAN-2001,	2001MO-US01354
XX		
PR	31-JAN-2000,	2000US-0179065
PR	04-FEB-2000,	2000US-0180628
PR	24-FEB-2000,	2000US-0184664
PR	02-MAR-2000,	2000US-0186509
PR	16-MAR-2000,	2000US-0189874
PR	17-MAR-2000,	2000US-0190077
PR	18-APR-2000,	2000US-01968123
PR	19-MAY-2000,	2000US-0205515
PR	07-JUN-2000,	2000US-0209467
PR	28-JUN-2000,	2000US-0214886
PR	30-JUL-2000,	2000US-0215133
PR	07-JUL-2000,	2000US-0216647
PR	11-JUL-2000,	2000US-0216880
PR	11-JUL-2000,	2000US-0217487
PR	14-JUL-2000,	2000US-0218290
PR	26-JUL-2000,	2000US-0220966
PR	14-AUG-2000,	2000US-0224518
PR	14-AUG-2000,	2000US-0224519
PR	14-AUG-2000,	2000US-0225213
PR	14-AUG-2000,	2000US-0225214
PR	14-AUG-2000,	2000US-0225266
PR	14-AUG-2000,	2000US-0225267
PR	14-AUG-2000,	2000US-0225268
PR	14-AUG-2000,	2000US-0225270
PR	14-AUG-2000,	2000US-0225447
PR	14-AUG-2000,	2000US-0225757
PR	14-AUG-2000,	2000US-0225758
PR	14-AUG-2000,	2000US-0225759
PR	18-AUG-2000,	2000US-0226279
PR	22-AUG-2000,	2000US-0226681
PR	22-AUG-2000,	2000US-0226882
PR	23-AUG-2000,	2000US-0227182
PR	30-AUG-2000,	2000US-0227009
PR	01-SEP-2000,	2000US-0228924
PR	01-SEP-2000,	2000US-0229287
PR	01-SEP-2000,	2000US-0229343
PR	01-SEP-2000,	2000US-0229344
PR	01-SEP-2000,	2000US-0229345
PR	05-SEP-2000,	2000US-0229350
PR	05-SEP-2000,	2000US-0229517
PR	06-SEP-2000,	2000US-0230433
PR	06-SEP-2000,	2000US-0230436
PR	08-SEP-2000,	2000US-0231142
PR	08-SEP-2000,	2000US-0231143
PR	08-SEP-2000,	2000US-0231144
PR	08-SEP-2000,	2000US-0231145
PR	08-SEP-2000,	2000US-0231146
PR	08-SEP-2000,	2000US-0231147
PR	08-SEP-2000,	2000US-0231148
PR	08-SEP-2000,	2000US-0231149
PR	08-SEP-2000,	2000US-0231150
PR	08-SEP-2000,	2000US-0231151
PR	08-SEP-2000,	2000US-0231152
PR	08-SEP-2000,	2000US-0231153
PR	08-SEP-2000,	2000US-0231154
PR	08-SEP-2000,	2000US-0231155
PR	08-SEP-2000,	2000US-0231156
PR	08-SEP-2000,	2000US-0231157
PR	08-SEP-2000,	2000US-0231158
PR	08-SEP-2000,	2000US-0231159
PR	08-SEP-2000,	2000US-0231160
PR	08-SEP-2000,	2000US-0231161
PR	08-SEP-2000,	2000US-0231162
PR	08-SEP-2000,	2000US-0231163
PR	08-SEP-2000,	2000US-0231164
PR	08-SEP-2000,	2000US-0231165
PR	08-SEP-2000,	2000US-0231166
PR	08-SEP-2000,	2000US-0231167
PR	08-SEP-2000,	2000US-0231168
PR	08-SEP-2000,	2000US-0231169
PR	08-SEP-2000,	2000US-0231170
PR	08-SEP-2000,	2000US-0231171
PR	08-SEP-2000,	2000US-0231172
PR	08-SEP-2000,	2000US-0231173
PR	08-SEP-2000,	2000US-0231174
PR	08-SEP-2000,	2000US-0231175
PR	08-SEP-2000,	2000US-0231176
PR	08-SEP-2000,	2000US-0231177
PR	08-SEP-2000,	2000US-0231178
PR	08-SEP-2000,	2000US-0231179
PR	08-SEP-2000,	2000US-0231180
PR	08-SEP-2000,	2000US-0231181
PR	08-SEP-2000,	2000US-0231182
PR	08-SEP-2000,	2000US-0231183
PR	08-SEP-2000,	2000US-0231184
PR	08-SEP-2000,	2000US-0231185
PR	08-SEP-2000,	2000US-0231186
PR	08-SEP-2000,	2000US-0231187
PR	08-SEP-2000,	2000US-0231188
PR	08-SEP-2000,	2000US-0231189
PR	08-SEP-2000,	2000US-0231190
PR	08-SEP-2000,	

PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	13-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	01-NOV-2000	2000US-0241826
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246609
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	17-NOV-2000	2000US-0246613
PR	17-NOV-2000	2000US-0246920
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249212
PR	17-NOV-2000	2000US-0249213
PR	17-NOV-2000	2000US-0249214
PR	17-NOV-2000	2000US-0249215
PR	17-NOV-2000	2000US-0249216
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249218
PR	17-NOV-2000	2000US-0249224
PR	17-NOV-2000	2000US-0249243
PR	17-NOV-2000	2000US-0249244
PR	17-NOV-2000	2000US-0249245
PR	17-NOV-2000	2000US-0249246
PR	17-NOV-2000	2000US-0249247
PR	17-NOV-2000	2000US-0249249
PR	17-NOV-2000	2000US-0249297
PR	17-NOV-2000	2000US-0249300
PR	01-DEC-2000	2000US-0250160
PR	01-DEC-2000	2000US-0250361
PR	05-DEC-2000	2000US-0250391
PR	05-DEC-2000	2000US-0250392
PR	05-DEC-2000	2000US-0250393
PR	05-DEC-2000	2000US-0251198
PR	06-DEC-2000	2000US-0256719
PR	08-DEC-2000	2000US-0251475
PR	08-DEC-2000	2000US-0251476
PR	08-DEC-2000	2000US-0251477
PR	08-DEC-2000	2000US-0251478
PR	08-DEC-2000	2000US-0251479
PR	08-DEC-2000	2000US-0251480
PR	08-DEC-2000	2000US-0251481
PR	08-DEC-2000	2000US-0251482
PR	08-DEC-2000	2000US-0251483
PR	08-DEC-2000	2000US-0251484
PR	08-DEC-2000	2000US-0251485
PR	08-DEC-2000	2000US-0251486
PR	08-DEC-2000	2000US-0251487
PR	08-DEC-2000	2000US-0251488
PR	08-DEC-2000	2000US-0251489
PR	08-DEC-2000	2000US-0251490
PR	08-DEC-2000	2000US-0251491
PR	08-DEC-2000	2000US-0251492
PR	08-DEC-2000	2000US-0251493
PR	08-DEC-2000	2000US-0251494
PR	08-DEC-2000	2000US-0251495
PR	08-DEC-2000	2000US-0251496
PR	08-DEC-2000	2000US-0251497
PR	08-DEC-2000	2000US-0251498
PR	08-DEC-2000	2000US-0251499
PR	08-DEC-2000	2000US-0251500
PR	08-DEC-2000	2000US-0251501
PR	08-DEC-2000	2000US-0251502
PR	08-DEC-2000	2000US-0251503
PR	08-DEC-2000	2000US-0251504
PR	08-DEC-2000	2000US-0251505
PR	08-DEC-2000	2000US-0251506
PR	08-DEC-2000	2000US-0251507
PR	08-DEC-2000	

DR WPI: 2001-483426/52.
 DR N-PSDB: AAK57561.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PS Claim 11; SEQ ID NO 12373; 3071bp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (II) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic antigen genomic
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX Sequence 85 AA:
 SQ
 Query Match 13.1%; Score 62; DB 22; Length 85;
 Best Local Similarity 37.7%; Pred. No. 30; Indels 18; Gaps 3;
 Matches 23; Conservative 2; Mismatches 18; Gaps 3;
 OY 1 PPAAPRPGSRARGNSSNHLGYCAEAPADPSTLYRDLPAEDSRGROGD-----AP 55
 Db 1 PPRY-TLDEGLKG-----AGLEGAGRSALRRRGPSADGREGTPVAGHCAP 47
 OY 56 T 56
 Db 48 T 48
 RESULT 12
 ID AAK57626 standard; Protein; 51 AA.
 AC AAK57626;
 AC AAK57626;
 13-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #18522.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 PD 01-NOV-2001.
 PF 20-APR-2001; 2001WO-US12865.
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208641P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71.
 DR N-PSDB: AAS59585.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 PS Example 1; SEQ ID NO 18821; 1069pp; English.
 XX
 CC Sequences AAK39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 51 AA:
 SQ
 Query Match 12.8%; Score 61; DB 22; Length 51;
 Best Local Similarity 38.6%; Pred. No. 22;
 Matches 22; Conservative 4; Mismatches 21; Indels 10; Gaps 4;
 OY 7 RFGSRARGNSSNHLGYCAEAPADPSTLYRDLPAEDSRGROGDAPTEDD-YWG 62
 Db 1 RP-HRRRTNTTS----VAELRNPSOPTT--HASGETCGRGCGENHAGGTCTYWG 48
 RESULT 13
 ID AAK50651 standard; Protein; 91 AA.
 AC AAK50651;
 AC AAK50651;
 13-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #11547.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 PD 01-NOV-2001.
 PF 20-APR-2001; 2001WO-US12865.
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208641P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

```

XX DR      WPI: 2001-616774/71.
XX DR      N-PSDB: AAS59549.
XX PT      Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT      vaccinating against and diagnosing infections, especially useful for
XX PT      treating acne vulgaris.
XX PS      Example 1: SEQ ID NO 11846; 1069pp; English.
XX CC      Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC      polypeptides. The proteins and their associated DNA sequences are used in
XX CC      the treatment, prevention and diagnosis of medical conditions caused by
XX CC      P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC      pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX CC      P. acnes is also involved in infections of bone, joints and the central
XX CC      nervous system, however it is particularly involved in the inflammatory
XX CC      lesions associated with acne vulgaris. A method for detecting the
XX CC      presence or absence of P. acnes in a patient comprises contacting a
XX CC      sample with a binding agent that binds to the proteins of the invention
XX CC      and determining the amount of bound protein in the sample. The
XX CC      polypeptides may be used as antigens in the production of antibodies
XX CC      specific for P. acnes proteins. These antibodies can be used to
XX CC      downregulate expression and activity of P. acnes polypeptides and
XX CC      therefore treat P. acnes infections. The antibodies may also be used as
XX CC      diagnostic agents for determining P. acnes presence, for example, by
XX CC      enzyme linked immunosorbent assay (ELISA).
XX CC      Note: The sequence data for this patent did not form part of the printed
XX CC      specification, but was obtained in electronic format directly from WIPO
XX CC      at ftp.wipo.int/pub/published_pct_sequences.
XX SQ      Sequence 91 AA:

Query Match      12.8%: Score 61; DB 22; Length 91;
Best Local Similarity 32.6%: Pred. No. 41;
Matches 14; Conservative 5; Mismatches 14; Indels 10; Gaps 1;

OY 40 DLPADSNGRGCD-----APTEDDYGCYGGEDGRGE 72
Db 21 DLPARSEBGERGLDLMGRGLGHVDAANSSDNMODHGDGDETK 63

RESULT 14
AAO08818
ID AAO08818 standard; Protein; 99 AA.
AC AAO08818:
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 22710.
XX
XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX PF
XX 28-FEB-2000; 2000US-0515126.
XX PR
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-514838/56.

```

```

XX DR      N-PSDB: AAI88749.
XX PT      Isolated nucleic acids and polypeptides, useful for preventing
XX PT      diagnosing and treating e.g. leukemia, inflammation and immune
XX PT      disorders.
XX PS      Claim 20: SEQ ID NO 22710; 1399pp + Sequence Listing; English.
XX CC      The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC      the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX CC      cytokine, cell proliferation or cell differentiation or which may induce
XX CC      production of other cytokines in other cell populations. The
XX CC      polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC      peptide therapy. The polypeptides have various cytokine-like activities,
XX CC      e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC      activity, tissue growth factor activity, immunomodulatory activity and
XX CC      activin/inhibin activity and may be useful in the diagnosis and/or
XX CC      treatment of cancer, leukemia, nervous system disorders, arthritis and
XX CC      inflammation.
XX CC      Note: The sequence data for this patent did not form part of the printed
XX CC      specification, but was obtained in electronic format directly from WIPO
XX CC      at ftp.wipo.int/pub/published_pct_sequences.
XX SQ      Sequence 99 AA:

Query Match      12.5%: Score 59.5; DB 22; Length 99;
Best Local Similarity 30.3%: Pred. No. 66;
Matches 20; Conservative 4; Mismatches 31; Indels 11; Gaps 3;

OY 1 PPAAPTRP---GSRANGSSNNHLYGVAE--AGAPAD-----PSTLYRDLPAEDSRGR 49
Db 9 PPFKPGPPQKRGKGAPKPGCEHFFWVLEKSTGFPKTKRYVRYSPRYVHPPAPQKEGF 68
OY 50 QGGDAP 55
Db 69 PGGSPP 74

RESULT 15
AAU59903
ID AAU59903 standard; Protein; 69 AA.
AC AAU59903:
XX
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #20799.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX PF
XX 21-APR-2000; 2000US-199047P.
XX PR
XX 02-JUN-2000; 2000US-208841P.
XX PR
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI: 2001-616774/71.
XX
XX N-PSDB: AAS59607.
XX

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 13.333 Seconds
(without alignments)
110.336 Million cell updates/sec

Title: US-09-972-546-2_COPY_261_310

Perfect score: 284
Sequence: 1 NPMACDCRARRPLMWFORAR.....PPERGRDLRALREADFGAC 50

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters: 23178

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	16.9	97	4	US-08-973-544-4 Sequence 4, Appl1
2	47	16.5	74	1	US-08-379-538-2 Sequence 2, Appl1
3	47	16.5	74	1	US-08-379-538-3 Sequence 3, Appl1
4	47	16.5	74	1	US-08-379-538-5 Sequence 5, Appl1
5	46	16.2	74	1	US-08-379-538-4 Sequence 4, Appl1
6	44	15.5	53	3	US-08-688-871-63 Sequence 63, Appl1
7	43.5	15.3	93	4	US-09-069-023-15 Sequence 15, Appl1
8	42.5	15.0	41	4	US-09-227-357-618 Sequence 618, App
9	42.5	15.0	53	2	US-08-799-173A-13 Sequence 13, Appl1
10	42.5	15.0	55	1	US-07-863-021B-20 Sequence 20, Appl1
11	42.5	15.0	55	5	PCT-US93-03164-20 Sequence 20, Appl1
12	42.5	15.0	97	4	US-08-905-223-444 Sequence 444, App
13	41	14.4	49	4	US-09-208-140-10 Sequence 10, Appl1
14	41	14.4	60	4	US-08-444-818-136 Sequence 136, App
15	41	14.4	99	2	US-09-047-125-21 Sequence 21, Appl1
16	41	14.4	99	3	US-07-736-335E-21 Sequence 21, Appl1
17	40.5	14.3	53	1	US-07-862-021B-18 Sequence 18, Appl1
18	40.5	14.3	53	5	PCT-US93-03164-18 Sequence 18, Appl1
19	40.5	14.3	66	2	US-08-836-620A-11 Sequence 11, Appl1
20	40.5	14.3	66	2	US-08-836-620A-12 Sequence 12, Appl1
21	40	14.1	49	4	US-09-208-140-9 Sequence 9, Appl1
22	40	14.1	49	4	US-09-208-140-11 Sequence 11, Appl1
23	40	14.1	50	4	US-09-208-140-8 Sequence 8, Appl1
24	39.5	13.9	49	2	US-08-726-306A-149 Sequence 149, App
25	39.5	13.9	56	1	US-08-538-165A-7 Sequence 7, Appl1
26	39.5	13.9	99	2	US-08-437-607A-43 Sequence 43, Appl1
27	39	13.7	61	4	US-09-135-020-108 Sequence 108, App

28	39	13.7	61	4	US-09-135-010A-108 Sequence 108, App
29	39	13.7	61	4	US-09-444-871-108 Sequence 108, App
30	39	13.7	61	4	US-09-597-735-108 Sequence 108, App
31	39	13.7	61	4	US-09-444-295-108 Sequence 108, App
32	39	13.7	61	4	US-09-597-732-108 Sequence 108, App
33	39	13.7	80	2	US-08-726-306A-173 Sequence 173, App
34	39	13.7	89	4	US-09-087-465-20 Sequence 20, Appl1
35	38.5	13.6	56	1	US-08-538-165A-1 Sequence 1, Appl1
36	38.5	13.6	62	2	US-08-687-702-19 Sequence 19, Appl1
37	38.5	13.6	63	4	US-09-099-631A-12 Sequence 12, Appl1
38	38.5	13.6	77	4	US-08-858-207A-285 Sequence 285, App
39	38	13.4	42	1	US-08-262-037-42 Sequence 42, Appl1
40	38	13.4	50	1	US-08-262-037-43 Sequence 21, Appl1
41	38	13.4	68	2	US-08-484-434C-21 Sequence 43, Appl1
42	38	13.4	77	4	US-09-605-785-558 Sequence 558, App
43	38	13.4	91	1	US-08-469-202-18 Sequence 18, Appl1
44	37.5	13.2	43	1	US-08-050-319B-43 Sequence 43, Appl1
45	37.5	13.2	43	2	US-08-465-982-43 Sequence 43, Appl1

ALIGNMENTS

RESULT 1
US-08-973-544-4
; Sequence 4, Application US/08973544
; Patent No. 6338950
GENERAL INFORMATION:
APPLICANT: WEISS, Elisabeth
TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,544
FILING DATE: 18-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT EP 96/02663
FILING DATE: 20-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95109511.6
FILING DATE: 20-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95112201.9
FILING DATE: 03-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kilts, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P8341-7073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-5000
TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-973-544-4
Query Match 16.9%; Score 48; DB 4; Length 97;

Best Local Similarity 36.5%; Pred. No. 27;
Matches 19; Conservative 7; Mismatches 16; Indels 10; Gaps 4;

```
QY      3 WACDCRRLP-WAMFORARVSSDYTCATPPERGRDLRALRE---ADFQAC 50
      11 : : : : : : : : : : : : : : : : : : : : : : : : : :
      45 WAGGSEDELHVASLQRLPLVPSE-----GPDLRGDKRKGTEKEDPRADY-AC 90
Db
```

RESULT 2

Sequence 2, Application US/08379538
Patent No. 5804554
GENERAL INFORMATION:
APPLICANT: Volkmann, Robert A.
APPLICANT: Saccamano, Nicholas A.
APPLICANT: Nason II, Deane M.
APPLICANT: Heck, Steven D.
APPLICANT: Ronau, Robert T.
TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES
TITLE OF INVENTION: FROM FILISTATA HIBERNALIS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pfizer Inc
STREET: 235 East 42nd Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/379,538
FILING DATE: 3-MAY-1995
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/887073
FILING DATE: 21-MAY-1992
APPLICATION NUMBER: PCI/US93/03921
FILING DATE: 30-APRIL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Zielinski, Bryan
REGISTRATION NUMBER: 34,462
REFERENCE/DOCKET NUMBER: PC8175A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 573-4585
TELEFAX: (212) 573-1939
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 74 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Filistata hibernalis
TISSUE TYPE: venom

Query Match	16.5%	Score 47;	DB 1;	Length 74;
Best Local Similarity	30.0%;	Pred. No. 27;		
Matches	9;	Conservative	5;	Mismatches 12;
				Indels 4;
				Gaps 1.

```

QY      5 CDCRARPLWAFQARARVSSSDVTCTATPPER 34
          | | : | | : | | : |
DB      49 CKCD---WSWSQRIDDMRADYSCKCPEDQ 74

```

RESULT 3

US-08-379-538-3
; sequence 3, Application US/08379538

GENERAL INFORMATION:
 APPLICANT: Volkmann, Robert A.
 APPLICANT: Saccomano, Nicholas A.
 APPLICANT: Nason II, Deane M.
 APPLICANT: Heck, Steven D.
 APPLICANT: Ronau, Robert T.
 TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES
 TITLE OF INVENTION: FROM FILISTATA HIBERNALIS
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pfizer Inc
 STREET: 235 East 42nd Street
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10017
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/379,558
 FILING DATE: 3-MAY-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/887073
 FILING DATE: 21-MAY-1992
 APPLICATION NUMBER: PCT/US93/03921
 FILING DATE: 30-APRIL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Zielinski, Bryan
 REGISTRATION NUMBER: 34,462
 REFERENCE/DOCKET NUMBER: PC8175A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 573-4585
 TELEFAX: (212) 573-1939
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 74 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Filistata hibernalis
 TISSUE TYPE: venom

Query Match	16.5%;	Score 47;	DB 1;	Length 74;
Best Local Similarity	30.0%;	Pred. No. 27;		
Matches	9;	Conservative	5;	Mismatches 12;
				Indels 4;
				Gaps 1.

```

QY      5 CDCRRLPMWFQRRARVSSSDVTCAITPER 34
          | | | | | : | | | : :
Db 49 CKCG---WSWSQRIDDRADYSCKCPEDQ 74

```

RESULT 4

US-08-379-538-5
; Sequence 5, Application US/08379538
; Patent No. 5804554

GENERAL INFORMATION:
APPLICANT: Volkmann, Robert A.
APPLICANT: Saccomano, Nicholas A.
APPLICANT: Nason II, Deane M.
APPLICANT: Heck, Steven D.
APPLICANT: Konau, Robert T.

```

1  TITLE OF INVENTION:  CALCIUM CHANNEL BLOCKING POLYPEPTIDES
2  TITLE OF INVENTION:  FROM FILISTATA HIBERNALIS
3  NUMBER OF SEQUENCES:  7
4  CORRESPONDENCE ADDRESSES:
5  ADDRESSEE:  Pfizer Inc
6  STREET:  235 East 42nd Street
7  CITY:  New York
8  STATE:  New York
9  COUNTRY:  USA
10 ZIP:  10017
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  Patentin Release #1.0, Version #1.25
17
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER:  US/08/379,538
20 FILING DATE:  3-MAY-1995
21 CLASSIFICATION:  514
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER:  US 07/887073
24 FILING DATE:  21-MAY-1992
25 APPLICATION NUMBER:  PCT/US93/03921
26 FILING DATE:  30-APRIL-1993
27 ATTORNEY/AGENT INFORMATION:
28 NAME:  Zieginski, Bryan
29 REGISTRATION NUMBER:  34,462
30 REFERENCE/DOCKET NUMBER:  PC8175A
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE:  (212) 573-4585
33 TELEFAX:  (212) 573-1939
34 INFORMATION FOR SEQ ID NO:  5:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH:  74 amino acids
37 TYPE:  amino acid
38 STRANDEDNESS:  single
39 TOPOLOGY:  linear
40 MOLECULE TYPE:  protein
41 HYPOTHEICAL:  NO
42 ANTI-SENSE:  NO
43 ORIGINAL SOURCE:
44 ORGANISM:  Filistata hibernalis
45 TISSUE TYPE:  venom
46
47 US-08-379-538-5
48
49 Query Match 16.5%, Score 47; DB 1; Length 74;
50 Best Local Similarity 30.0%; Pred. No. 27;
51 Matches 9; Conservative 5; Mismatches 12; Indels 4; Gaps 1.
52
53 5 CDCRARPLAMVFORARVSSSDVTCATPPER 34
54 | | | | | | | | | | | | | | | |
55 Db 49 KCCK---WMSQRIIDWRADYSCCKPEQD 74
56
57 RESULT 5
58 US-08-379-538-4
59 Sequence 4, Application US/08379538
60 Patent No. 5804554
61 GENERAL INFORMATION:
62 APPLICANT:  Volkmann, Robert A.
63 APPLICANT:  Saccomano, Nicholas A.
64 APPLICANT:  Nason II, Deane M.
65 APPLICANT:  Heck, Steven D.
66 APPLICANT:  Ronau, Robert T.
67 TITLE OF INVENTION:  CALCIUM CHANNEL BLOCKING POLYPEPTIDES
68 TITLE OF INVENTION:  FROM FILISTATA HIBERNALIS
69 NUMBER OF SEQUENCES:  7
70 CORRESPONDENCE ADDRESSES:
71 ADDRESSEE:  Pfizer Inc
72 STREET:  235 East 42nd Street
73 CITY:  New York
74 STATE:  New York
75 COUNTRY:  USA

```

```

1  ZIP: 10017
2  COMPUTER READABLE FORM:
3  MEDIUM TYPE: Floppy disk
4  COMPUTER: IBM PC compatible
5  OPERATING SYSTEM: PC-DOS/MS-DOS
6  SOFTWARE: Patentin Release #1.0, Version #1.25
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/08/379,538
9  FILING DATE: 3-MAY-1995
10 CLASSIFICATION: 514
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 07/887073
13 FILING DATE: 21-MAY-1992
14 APPLICATION NUMBER: PCT/US93/03921
15 FILING DATE: 30-APRIL-1993
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Zielinski, Bryan
18 REGISTRATION NUMBER: 34,462
19 REFERENCE/DOCKET NUMBER: PC8175A
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (212) 573-4585
22 TELEFAX: (212) 573-1939
23 INFORMATION FOR SEQ ID NO: 4:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 74 amino acids
26 TYPE: amino acid
27 STRANDEDNESS: single
28 TOPOLOGY: linear
29 MOLECULE TYPE: protein
30 HYPOTHEetical: NO
31 ANTI-SENSE: NO
32 ORIGINAL SOURCE:
33 ORGANISM: Filistata hibernalis
34 TISSUE TYPE: venom
35 US-08-379-538-4
36
37 Query Match 16.2% Score 46; DB 1; Length 74;
38 Best Local Similarity 30.0%; Pred. No. 36;
39 Matches 9; Conservative 4; Mismatches 13; Indels 4; Gaps 1;
40
41 OY 5 CDCRRAPLAWFORARVSSDVTCTPPER 34
42 | | | | | : | | | | :
43 Db 49 CKCG---WSMSQRITDWRADYCNCKPEDQ 74
44
45 RESULT 6
46 US-08-685-871-63
47 Sequence 63, Application US/08685871
48 Patent No. 6013499
49 GENERAL INFORMATION:
50 APPLICANT: NARUKIYA, Shuh
51 APPLICANT: IMAMATSU, Akihito
52 TITLE OF INVENTION: RHO TARGET PROTEIN KINASE P160
53 NUMBER OF SEQUENCES: 68
54 CORRESPONDENCE ADDRESS:
55 ADDRESSEE: Foley & Lardner
56 STREET: 3000 K Street, N.W., Suite 500
57 CITY: Washington
58 STATE: D.C.
59 COUNTRY: USA
60 ZIP: 20007-5109
61 COMPUTER READABLE FORM:
62 MEDIUM TYPE: Floppy disk
63 COMPUTER: IBM PC compatible
64 OPERATING SYSTEM: PC-DOS/MS-DOS
65 SOFTWARE: Patentin Release #1.0, Version #1.30
66 CURRENT APPLICATION DATA:
67 APPLICATION NUMBER: US/08/685,871
68 FILING DATE: 24-JUL-1996
69 CLASSIFICATION: 435
70 PRIOR APPLICATION DATA:
71 APPLICATION NUMBER: JP 8-184102
72 FILING DATE: 25-JUN-1996

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-262553
FILING DATE: 14-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16887/845
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-685-871-63

Query Match          15.5%; Score 44; DB 3; Length 53;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 2 PWACDRAPLWAFQ 17
Db 11 PANCDAKPLMHVK 26

RESULT 7
US-09-069-023-15
Sequence 15, Application US/09069023A
Patent No. 6348573
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohiro
APPLICANT: Koseki, Takeyoshi
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
FILE REFERENCE: UM-03333
CURRENT APPLICATION NUMBER: US/09/069,023A
CURRENT FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 93
TYPE: PRT
ORGANISM: Homo sapiens
US-09-069-023-15

Query Match          15.3%; Score 43.5; DB 4; Length 93;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 12; Conservative 2; Mismatches 14; Indels 5; Gaps 1;

Oy 23 SSDVTCATPPEROG----RDLRALREADFOAC 50
Db 5 SQHVPTCSPPKRSGLARVTFDLYLKPKDFICG 37

RESULT 8
US-09-227-357-618
Sequence 618, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
```

```

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 618
LENGTH: 41
TYPE: PRT
ORGANISM: Homo sapiens
US-09-227-357-618

Query Match          15.0%; Score 42.5; DB 4; Length 41;
Best Local Similarity 27.3%; Pred. No. 53;
Matches 12; Conservative 2; Mismatches 13; Indels 17; Gaps 2;
```


QY 13 MAMFORARVSSDVTCATPEROGRDLRALREADFOAC 50
DB 5 MAMSORRH-----PGRPKDQGRGLMTQSRSGDARCC 37

RESULT 9
US-08-799-173A-13

Sequence 13, Application US/08799173A
Patent No. 5871969
GENERAL INFORMATION:
APPLICANT: HASTINGS, GREGG,
APPLICANT: PATRICK J. DILLON
TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/799,173A
FILING DATE: 11-FEB-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-799-173A-13

Query Match 15.0% Score 42.5; DB 2; Length 53;
Best Local Similarity 22.7%; Pred. No. 71;
Matches 10; Conservative 8; Mismatches 21; Indels 5; Gaps 1;

QY 7 CRARPLMAMFORARVSSDVTCATPEROGRDLRALREADFOAC 50
DB 1 CRMRPWTAMSECTKLCGGCI-----QERYMTVKRFRKSSQFTSC 39

RESULT 10
US-07-862-021B-20

Sequence 20, Application US/07862021B
Patent No. 5279966
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M
APPLICANT: Klar, Avihu
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,021B
FILING DATE: 19920405
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-862-021B-20

Query Match 15.0% Score 42.5; DB 1; Length 55;
Best Local Similarity 22.7%; Pred. No. 74;
Matches 10; Conservative 8; Mismatches 21; Indels 5; Gaps 1;

QY 7 CRARPLMAMFORARVSSDVTCATPEROGRDLRALREADFOAC 50
DB 3 CRMRPWTAMSECTKLCGGCI-----QERYMTVKRFRKSSQFTSC 41

RESULT 11
PCT-US93-03164-20

Sequence 20, Application PC/TUS9303164
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M
APPLICANT: Klar, Avihu
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03164
FILING DATE: 19930402
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids

TYPE: AMINO ACID
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-03164-20

Query Match 15.0%; Score 42.5; DB 5; Length 55;
Best Local Similarity 22.7%; Pred. No. 74;
Matches 10; Conservative 8; Mismatches 21; Indels 5; Gaps 1;

OY 7 CRARPLMAMFORARVSSSDVTCAPPEFQGDRLALRPADQAC 50
DB 3 CRMRPWTAMSECTKLCGGGI-----QERYMTVKRKFKSSOPTSC 41

RESULT 12
US-08-905-223-444
Sequence 444, Application US/08905223
Patent No. 6222029

GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-8550
INFORMATION FOR SEQ ID NO: 444:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig-peptide
LOCATION: -35--1
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 6.4
OTHER INFORMATION: seq SACLILCPWTNP/OL

US-08-905-223-444

Query Match 15.0%; Score 42.5; DB 4; Length 97;
Best Local Similarity 37.0%; Pred. No. 1.4e+02;
Matches 10; Conservative 5; Mismatches 9; Indels 3; Gaps 1;

OY 10 RPLMAM---FORARVSSSDVTCAPPE 33
DB 62 RPSXKMSLFPATVGSSTVSAVTPD 88

RESULT 13
US-09-208-140-10
Sequence 10, Application US/09208140
Patent No. 6228576

GENERAL INFORMATION:
APPLICANT: Del Vecchio, Alfred
TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN
TITLE OF INVENTION: AND METHODS THEREOF TO IDENTIFY ANTIVIRAL COMPOUNDS
FILE REFERENCE: P50743
CURRENT FILING DATE: 1998-12-09
CURRENT APPLICATION NUMBER: US/09/208,140
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 49
TYPE: PRT
ORGANISM: Viral

US-09-208-140-10

Query Match 14.4%; Score 41; DB 4; Length 49;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 9 ARPLMAMF 16
DB 25 ARPLMAMF 32

RESULT 14
US-08-444-818-136
Sequence 136, Application US/08444818
Patent No. 6150867

GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisha A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-444-818-136

Query Match 14.4%; Score 41; DB 4; Length 60;

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 38.8889 seconds

(without alignments)
264.918 Million cell updates/sec

Title: US-09-972-546-2_COPY_261_310

Sequence: 1 NPMACDCRRAPLMWFORAR.....PEPGCRDRLALREADFOAC 50

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Minimum DB seq length: 40
Maximum DB seq length: 100
Listing first 45 summaries 101114

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP virus: *
16: SP bacteriophage: *
17: SP archaea: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	20.1	100	12	064884 mouse adeno
2	52	18.3	91	16	09K3B9 streptomyc
3	50.5	17.8	67	12	098283 molluscum c
4	50	17.6	85	4	013659
5	49.5	17.4	88	16	093DB8
6	47.5	16.7	91	4	000452
7	47	16.5	91	2	091UN6
8	47	16.5	91	2	007301
9	46.5	16.4	53	4	000454
10	46.5	16.4	80	5	09N9V9
11	46.5	16.4	96	6	09BGP8
12	46.5	16.4	97	4	000453
13	46.5	16.4	99	2	033275
14	45	15.8	62	16	09PBF8
15	45	15.8	73	15	087603
16	45	15.8	91	2	09RCE4

17	45	15.8	92	2	005608
18	45	15.8	94	16	09K493
19	45	15.8	100	11	099N01
20	44	15.5	67	5	023727
21	44	15.5	84	5	016008
22	44	15.5	90	5	044262
23	44	15.5	91	2	06047
24	44	15.5	91	2	06047
25	43.5	15.3	98	4	08TBX9
26	43.5	15.3	77	15	091685
27	43.5	15.3	82	6	046622
28	43.5	15.3	85	6	09GM76
29	43.5	15.3	91	16	08VKG3
30	43	15.1	69	10	094K36
31	43	15.1	70	5	023777
32	43	15.1	80	5	023774
33	43	15.1	86	4	09BVB7
34	43	15.1	86	16	09LE71
35	43	15.1	92	2	093604
36	42.5	15.0	49	16	098A12
37	42.5	15.0	74	15	098A12
38	42.5	15.0	80	16	098A12
39	42.5	15.0	80	16	098A12
40	42	14.8	53	16	08VIX2
41	42	14.8	65	3	096VG5
42	42	14.8	78	5	0967W3
43	42	14.8	79	16	09ASJ7
44	42	14.8	81	10	093VJ7
45	42	14.8	83	9	092X74

ALIGNMENTS

RESULT 1
ID 064884 PRELIMINARY: PRT: 100 AA.
AC 064884;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE E3 class 1 protein.
OS Mouse adenovirus type 1 (MAV-1).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90177241; PubMed=2137954;
RA Beard C.W., Ball A.O., Wooley E.H., Spindler K.R.;
RT "Transcription mapping of mouse adenovirus type 1 early region 3.";
RL Virology 175:81-90(1990).
DR EMBL: M31995; AAA42494.1; -;
SQ SEQUENCE 100 AA; 10908 MW; C8DDCC0E9C4BFA9 CRC64;

Query Match: 20.1%; Score 57; DB 12; Length 100;
Best Local Similarity 29.2%; Pred. No. 7.2;
Matches 14; Conservative 6; Mismatches 22; Indels 6; Gaps 1;
OY 2 PNMACDCRRAPLM-----AMFORARSSDYVCATPPRQGRDLRLR 43
Db 43 PSCQCGPASPWTNSSVTSPAOKTKWNSROYCPVPSSESTRGNAYR 90
RESULT 2
ID 09K3B9 PRELIMINARY: PRT: 91 AA.
AC 09K3B9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative redoxin.
GN NRDX OR SC05227 OR SC7E4.24C.
OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RX Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RA "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Borovok I., Karaffa L., Karaffa E., Cohen G., Aharonowitz Y.;
 RT "Cloned Streptomyces coelicolor A3(2) operon encoding ribonucleotide
 RT eukaryotic and viral ribonucleotide reductases.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M45;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL: AL359214; CAB94612.1;
 DR EMBL: AJ276618; CAB82484.1;
 SC SEQUENCE 91 AA; 9708 MW; 11980A14C0903163 CRC64;
 Query Match 18.3%; Score 52; DB 16; Length 91;
 Best Local Similarity 34.7%; Pred. No. 29;
 Matches 17; Conservative 3; Mismatches 29; Indels 0; Gaps 0;

QY 1 NPMACDCRARPPLMAMFORARVSSDYTCATPPERQGRDLRALREA 49
 Db 38 SPRRCQCARPPWPGCHDVRASSGPDVATPRALVSCPLPSLRAPCA 86

RESULT 3
 Q98283 PRELIMINARY; PRT; 67 AA.
 ID Q98283;
 AC Q98283;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MCL16R.
 GN MCL16R.
 OS Molluscum contagiosum virus subtype 1 (MCV1).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Molluscipoxvirus.
 OX NCBI_TaxID=10280;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=96325459; PubMed=8670425;
 RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
 RA Moss B.;
 RT "Genome sequence of a human tumorigenic poxvirus: Prediction of
 RT specific host response-evasion genes.";
 RL Science 273:813-816(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
 RA Moss B.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U60315; AAC55244.1;
 SC SEQUENCE 67 AA; 7693 MW; 962AA006ADDBAA2 CRC64;
 Query Match 17.8%; Score 50.5; DB 12; Length 67;
 Best Local Similarity 41.7%; Pred. No. 34;
 Matches 15; Conservative 4; Mismatches 10; Indels 7; Gaps 1;

QY 10 RPLMAMFORARVSSDYTCATPPERQGRDLRALREA 45
 Db 27 RRGARREFRRARASSAE-----RRARLELALRAA 55

RESULT 4
 Q13669 PRELIMINARY; PRT; 85 AA.
 ID Q13669;
 AC Q13669;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE Lst-1 protein.
 GN Lst-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=96006965; PubMed=7590964;
 RA Holzinger I., de Baey A., Wesser G., Kick G., Zwierzina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 RT TNF region.";
 RT Immunogenetics 42:315-322(1995).
 RL EMBL: U00921; AAB57724.1;
 SC SEQUENCE 85 AA; 9410 MW; 0276676D528C49A6 CRC64;
 Query Match 17.6%; Score 50; DB 4; Length 85;
 Best Local Similarity 25.6%; Pred. No. 49;
 Matches 20; Conservative 5; Mismatches 13; Indels 40; Gaps 5;

QY 3 WACDCRARPPLMAMFORARVSSDYTCATPPERQGRDLRALREA 35
 Db 11 WAGSC---FWQSFPCPPACVGCIEEAGSSRQELHYASLGRLPVPSSE-----GPDLR 61

QY 36 GRDLRALRE---ADFOAC 50
 Db 62 GRDKRGTKEDPRADY-AC 78

RESULT 5
 Q93DB8 PRELIMINARY; PRT; 88 AA.
 ID Q93DB8;
 AC Q93DB8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE AVHB7 (Agrobacterium virulence homologue VIRB7).
 GN AVHB7 OR ATU5168 OR AGR_PAT_227.
 OS Agrobacterium tumefaciens, and
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OG Plasmid AT.

Db 16 PVMATQTVLTVSGMTCSTCP 37

RESULT 8

ID 007301 PRELIMINARY; PRT: 91 AA.
AC 007301;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Mercutic transport protein periplasmic component precursor
DE (periplasmic mercury ion binding protein) (Mercury scavenger protein).
GN MERP.
OS Pseudomonas sp., and
OS Pseudomonas sp. ED23-33.
OG Plasmid pMR26.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306, 83781;
[1]

SEQUENCE FROM N.A.

SPECIES=Pseudomonas sp.; STRAIN-K-62;
MEDLINE=97311403; PubMed=9168120;
RA Kiyono M., Omura T., Inuzuka M., Fujimori H., Pan-Hou H.;
RT "Nucleotide sequence and expression of the organomercurial-resistance
determinants from a Pseudomonas K-62 plasmid pMR26.";
RT Gene 189:151-157(1997).
[2]

SEQUENCE FROM N.A.

SPECIES=Pseudomonas sp. ED23-33; TRANSPONSON-TN5058;
RA Minakina S., Minakina L., Kholodil G., Mindlin S., Gorlenko Z.H.,
RA Yurieva O., Nikiforov V.;
RT "Molecular inventory of transposons from environmental bacteria:
epidemic dissemination of Tn21-, Tn5041-, and Tn5053-related
transposons.";
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: D83080; BAA20336.1; -;
DR EMBL: Y17897; CAC14699.1; -;
DR HSSP: P04129; IATF.
DR InterPro: IPR001934; HeavyMe.transpt.
DR InterPro: IPR001802; HG_scarvenger.
DR Pfam: PF00403; HMA; 1.
DR PRINTS: PR00946; HGSCAVENGER.
DR PROSITE: PS01047; HMA_1; 1.
KW Plasmid; Signal.
FT SIGNAL 1 19
FT CHAIN 20 91
FT SEQUENCE 91 AA; 9475 MW; A9EBA62341075968 CRC64;
[1]

Query Match

Best Local Similarity 16.5%; Score 47; DB 2; Length 91;
Pred. No. 1.3e+02;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 11 PVMATQTVLTVSGMTCSTCP 32
Db 16 PVMATQTVLTVSGMTCATCP 37

RESULT 9

ID 000454 PRELIMINARY; PRT: 53 AA.
AC 000454;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE LST1 protein.
GN LST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.

RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RT differential regulation, initiation, and alternative splicing.";
RL Genomics 45:591-600(1997).
DR EMBL: AF000426; AAB87000.1; -;
SQ SEQUENCE 53 AA; 5887 MW; 5A9249680231E74F CRC64;

Query Match

Best Local Similarity 16.4%; Score 46.5; DB 4; Length 53;
Pred. No. 89;
Matches 10; Conservative 6; Mismatches 10; Indels 9; Gaps 3;

OY 13 WAMFORARVSSDVTCATPPEGRDILALRE--AFPOAC 50
Db 12 YASLQRLVPVPSSE-----GPDLRGRDKRGTEDEPRADY-AC 46

RESULT 10

ID 09N9V9 PRELIMINARY; PRT: 80 AA.
AC 09N9V9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Eukaryotic translation initiation factor 2 gamma (Fragment).
GN EIF2G.
OS Lithobius forficatus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
OX NCBI_TaxID=7552;
[1]
RP SEQUENCE FROM N.A.
MEDLINE=20519450; PubMed=11063691;
RA Kraus V., Reuter G.;
RT "Two genes become one: The genes encoding Heterochromatin protein
RT SU(VAR)3-9 and translation initiation factor subunit eIF2gamma are
RT joined to a dicistronic unit in holometabolic insects.";
RL Genetics 156:1157-1167(2000).
DR EMBL: AJ290958; CAB97508.1; -;
KW Initiation factor.
FT NON_TER 1 80
FT SEQUENCE 80 AA; 8902 MW; 9F5463C25BBFBF06 CRC64;
[1]

Query Match

Best Local Similarity 16.4%; Score 46.5; DB 5; Length 80;
Pred. No. 1.3e+02;
Matches 13; Conservative 3; Mismatches 13; Indels 3; Gaps 1;

OY 5 CDCRAPPPLNMFQARVSSDVTCATPPERQG 36
Db 38 CDCDACPRACVYSGSSKED--AFPCDRMG 66

RESULT 11

ID 09BGP8 PRELIMINARY; PRT: 96 AA.
AC 09BGP8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 9.9 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolpus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=FRONTAL LOBE LEFT;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA

RT Libraries: 1
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB056424; BAB33082.1;
 KW Hypothetical protein.
 SO SEQUENCE 96 AA; 9903 MW; 62E4A75153B2E38 CRC64;

Query Match
 Best Local Similarity 16.4%; Score 46.5; DB 6; Length 96;
 Matches 9; Conservative 4; Mismatches 7; Indels 3; Gaps 1;
 OY 5 CDCR---ARPLMAMFORARVSS 24
 DB 55 CSCRCGATAPVGAWLEKKNSS 77

RESULT 12
 ID 000453 PRELIMINARY: PRT: 97 AA.
 AC 000453;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DE 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 GN LST1 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Meier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing";
 RL Genomics 45:591-600(1997).
 DR EMBL: AF000425; AAB86899.1;
 SO SEQUENCE 97 AA; 10822 MW; AA03C761E787AF94 CRC64;

Query Match
 Best Local Similarity 16.4%; Score 46.5; DB 4; Length 97;
 Matches 16; Conservative 6; Mismatches 10; Indels 9; Gaps 3;

OY 13 WAMFORARVSSDYTCATPEPCGRDLRALRE--ADFQAC 50
 DB 56 YASIQRLPVPSS-----GPDLRGRDKRGTKEDPRADY-AC 90
 ID 033275 PRELIMINARY: PRT: 99 AA.
 AC 033275;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE 01-JAN-1998 (TREMBLrel. 05, Last annotation update)
 GN MCP protein.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RA Dasgupta N., Tyagi J.S.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y15717; CAA75732.1;
 SO SEQUENCE 99 AA; 10604 MW; 2768A2CD994D4D09 CRC64;

Query Match
 Best Local Similarity 16.4%; Score 46.5; DB 2; Length 99;
 Matches 11; Conservative 5; Mismatches 8; Indels 7; Gaps 1;

OY 7 CRARPL-----WAMFORARVSSDYTCAT 30
 DB 34 CKARPIAKSRMTWMSRRAIRSRVSTSS 64

RESULT 14
 ID 09PDF8 PRELIMINARY: PRT: 62 AA.
 AC 09PDF8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 GN Hypothetical protein Xf1421.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxId=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
 RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Facincani A.P., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Ferreira A.J.S., Ferreira V.C.A., Ferris J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitaajima J.P.,
 RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.C., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA de Oliveira R.R., Pereira G.A.G., Pereira H.A.Jr., Pasquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA da Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Trufel D., Tsai S.M., Tsubako M.H.,
 RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Zetubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 DR EMBL: AE003972; AAF84230.1;
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 62 AA; 7178 MW; 2E1427E4525F7CD2 CRC64;

Query Match
 Best Local Similarity 15.8%; Score 45; DB 16; Length 62;
 Matches 15; Conservative 5; Mismatches 20; Indels 4; Gaps 2;

OY 10 RPL-WAMFORARVSSDYTCATP---PERGCRDLRALREADFOA 49
 DB 7 RPIVGRWFARLACVSPHMTTDRPLQDPKQSRVATPEPFAA 50

RESULT 15
 ID 087603 PRELIMINARY: PRT: 73 AA.
 AC 087603;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 GN Rev protein (Fragment).
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).

Fri Apr 4 12:11:46 2003

us-09-972-546-2_261_310.szIm40-100.rspt

```
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P058;
RX MEDLINE=97138325; PubMed=8985351;
RA Bidollet-Ruche F., Brenques C., Galat-Luong A., Galat G., Pourrut X.,
  Vidal N., Veas F., Durand J.P., Cuny G.;
RT "Genetic diversity of simian immunodeficiency viruses from West
  African green monkeys: evidence of multiple genotypes within
  populations from the same geographical locale.";
RL J. Virol. 71:307-313(1997).
DR EMBL; U37210; AAC56159.1;
DR InterPro; IPR000625; REV_protein.
DR Pfam; PF00424; REV; 1.
FT NON_TER
SQ SEQUENCE 73 AA; 8217 MW; A31FDD937DC9BB4D CRC64;

Query Match 15.8%; Score 45; DB 15; Length 73;
Best Local Similarity 41.7%; Pred.No. 1.9e+02;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 8 RARPLWAFQARFVSSSDVTCATP 31
   ||||| :| :|
Db 14 RARPRWAHQROLIYLDRIATTP 37
```

Search completed: April 4, 2003, 08:28:01
Job time : 42.8889 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 12:07:27 ; Search time 75 Seconds
(without alignments)
746.203 Million cell updates/sec

Title: US-09-972-546-2
Perfect score: 2226
Sequence: 1 MLPGRRLRIQAPASACLLM.....LSAGLSPDLCLLLVPHHL 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq-101002:*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2226	100.0	420	23	AAO21477
2	2082	93.5	390	23	AAO21482
3	1977.5	88.8	807	22	ABG15173
4	911	40.9	461	23	AAO21478
5	887	39.8	421	23	AAO21483
6	836.5	37.6	241	22	ABG15171
7	811	36.4	392	23	AAO21487
8	808	36.3	472	21	AAAB33426
9	808	36.3	473	20	AAAY41745
10	808	36.3	473	21	AAAB44301

11	808	36.3	473	21	AAAB24410	Human PRO526 prote
12	808	36.3	473	21	AAAY93445	Human PRO526 anltic
13	808	36.3	473	22	AAU12362	Human PRO526 polyp
14	808	36.3	473	22	AAU04589	Human Nogo recepto
15	808	36.3	473	22	AAAB50908	Human PRO526 prote
16	808	36.3	473	22	AAAB59851	Human PRO526 prote
17	808	36.3	473	23	AAO21479	Human NGR1 protein
18	808	36.3	473	23	AAAB36555	Human PRO protein,
19	805	36.2	473	22	AAAB71057	Human secreted pro
20	805	36.2	473	22	AAAB71162	Human secreted pro
21	805	36.2	473	22	AAAB71164	Human secreted pro
22	804	36.1	473	22	AAAB71165	Human secreted pro
23	801	35.6	473	22	AAAB71166	Human secreted pro
24	792.5	35.6	473	22	AAAB71167	Human secreted pro
25	783.5	35.2	380	21	AAAY6244	Human mature TANCO
26	781.5	35.1	421	21	AAAB7101	Human mature TANCO
27	775	34.8	473	22	AAU04590	Human TANCO 393 ex
28	775	34.8	473	22	AAAB7102	Mouse Nogo recepto
29	775	34.8	473	22	AAAB7106	Mouse secreted pro
30	775	34.8	473	22	AAAB7108	Mouse secreted pro
31	772	34.7	473	22	AAAB7109	Mouse secreted pro
32	771	34.6	473	22	AAAB7167	Mouse secreted pro
33	771	34.6	473	22	AAAB7168	Mouse secreted pro
34	769.5	34.6	473	22	AAAB7169	Mouse secreted pro
35	767.5	34.5	447	22	AAAB7105	Mouse secreted pro
36	757.5	34.0	472	22	AAAB7116	Mouse secreted pro
37	757.5	28.6	423	22	AAAB7117	Mouse secreted pro
38	578	26.0	254	22	ABG15172	Mouse secreted pro
39	546	24.5	440	23	AAAB66666	Mouse mature TANCO
40	374.5	16.8	723	21	AAAB38400	Mouse mature TANCO
41	374	16.8	481	23	AAAB38400	Mouse mature TANCO
42	374	16.8	481	23	AAAB38400	Mouse mature TANCO
43	371	16.7	673	22	AAAB51131	Human novel polyep
44	369	16.6	611	21	AAAB46849	Human novel polyep
45	369	16.6	673	21	AAAB38323	Human novel polyep

ALIGNMENTS

RESULT 1	AAO21477	standard: Protein: 420 AA.
ID	AAO21477	
XX	AAO21477	
AC	AAO21477	
XX	AAO21477	
XX	AAO21477	
DT	15-AUG-2002	(first entry)
XX	15-AUG-2002	
DE	Human NGR2 protein sequence.	
XX	Human NGR2 protein sequence.	
KW	Cerebroprotective; neuroprotective; cytoskeletal; Nogo receptor homologue;	
KW	NGR2; NGR3; axonal stroke; central nervous system; CNS; cerebral injury;	
KW	spinal cord injury; stroke; demyelinating disease; multiple sclerosis;	
KW	monophasic demyelination; encephalomyelitis; Marchiafava-Bignami disease;	
KW	multifocal leukoencephalopathy; panencephalitis; Spongy degeneration;	
KW	Alexander's disease; Canavan's disease; metachromatic leukodystrophy;	
KW	Krabbe's disease; immune; bait protein; genetic mapping; gene therapy;	
XX	transgenic animal; unregulated cellular growth; cancer; tumour; human.	
OS	Homo sapiens.	
XX	Homo sapiens.	
XX	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	Peptide	1..30
FT	Region	/label= signal_peptide
FT	Region	31..59
FT	Region	/label= LRRNT_region
FT	Region	60..82
FT	Region	/label= LRR1_region
FT	Region	83..106
FT	Region	/label= LRR2_region
FT	Region	107..131
FT	Region	/label= LRR3_region
FT	Region	132..155

Db	1	MLPGRLRLQAPASACLLMLLALPLAAPSCPMCTCTCSSPPYVSCANNSSVPLSLPP	60
QY	61	STORLEFIONNNLIRLRBPFGSGNLLTTLTLEFSSNNLSTIYPGTFRHLQALBELDGDNRHL	120
Db	61	STORLEFIONNNLIRLRBPCTFGSNLLTTLTLEFSSNNLSTIYPGTFRHLQALBELDGDNRHL	120
QY	121	SLPDDTFOGGLERLQSLHLHYRCQLSLPGNIFRGIVSLQLYLQENSLLHQQDDLEFADLAN	180
Db	121	SLPDDTFOGGLERLQSLHLHYRCQLSLPGNIFRGIVSLQLYLQENSLLHQQDDLEFADLAN	180
QY	181	LSHLEFHNRLRLTLLEHVFNRGIGSLIDRLLLHGNRLQGVHRAAFRLSLTILLYFNNSLA	240
Db	181	LSHLEFHNRLRLTLLEHVFNRGIGSLIDRLLLHGNRLQGVHRAAFRLSLTILLYFNNSLA	240
QY	241	SILPGALADLPSELEFRLINANPWACDCARPLMAWFGARVSSSVTCATPPERQGLRL	300
Db	241	SILPGALADLPSELEFRLINANPWACDCARPLMAWFGARVSSSVTCATPPERQGLRL	300
QY	301	ALREDFQACPPAAPTFRGSRARGSSSNHLTYGVAEGAPPAAPSTLYRDLPAEDSRGRQ	360
Db	301	ALREDFQACPPAAPTFRGSRARGSSSNHLTYGVAEGAPPAAPSTLYRDLPAEDSRGRQ	360
QY	361	GSDAPTEDDYWGCGYGEGDQGEOMCPGAACQAPPPSRGALSAGLPSPILLCILLVPHLL	420
Db	361	GSDAPTEDDYWGCGYGEGDQGEOMCPGAACQAPPPSRGALSAGLPSPILLCILLVPHLL	420
RESULT 2			
AAO21482			
ID	AAO21482	standard. Protein: 390 AA.	
XX	AAO21482:		
AC			
XX			
DT	15-AUG-2002	(first entry)	
XX			
DE	Mature human Ngr2 protein sequence.		
XX			
XX	Cerebroprotective; neuroprotective; cytosstatic; Nogo receptor homologue;		
KW	Ngr2; NGR3; axonal growth; central nervous system; CNS; cerebral injury;		
KW	spinal cord injury; stroke; demyelinating disease; multiple sclerosis;		
KW	monophasic demyelination; encephalomyelitis; Marchiafava-Bignami disease;		
KW	multifocal leukoencephalopathy; panencephalitis; Spongy degeneration;		
KW	Alexander's disease; Canavan's disease; metachromatic leukodystrophy;		
KW	Krabbe's disease; immune; bait protein; gene therapy; gene mapping;		
KW	transgenic animal; unregulated cellular growth; cancer; tumour; human.		
XX			
OS	Homo sapiens.		
PN	WO200223059-A2.		
XX			
PD	11-APR-2002.		
XX			
XX	06-OCT-2001; 2001WO-US31488.		
XX			
XX	06-OCT-2000; 2000US-238361P.		
PA	(UYVA) UNIV YALE.		
PA	(BIOJ) BIOGEN INC.		
PI	Striltmatter SM, Cate RL, Sah DWY.		
XX			
DR	WPI; 2002-416677/44.		
XX			
PT	Novel Nogo receptor homolog poly peptide, Ngr2 or Ngr3, useful for		
PT	treating central nervous system disorder, cerebral injury, spinal cord		
PT	injury, stroke, and demyelinating diseases		
PS			
XX	Example 16; Page 106; 277pp; English.		
XX			
CC	The invention relates to a Nogo receptor homologue poly peptide, Ngr2 or		
CC	Ngr3, comprising a 50 amino acid LRRC sequence, a 284 amino acid NTR-RC		
CC	sequence, or a 420, 461 or 392 amino acid sequence, all given in the		
CC	specification. The Ngr3 protein or its binding antibody is useful for		

decreasing inhibition of axonal growth of a central nervous system (CNS) neuron, by contacting the neuron Ngr3 or its antibody, and for treating CNS disease, disorder or injury, Ngr3 or a vector comprising Ngr3 is useful for treating cerebral injury, spinal cord injury, stroke, demyelinating diseases, e.g. multiple sclerosis, monophasic demyelination, encephalomyelitis, multifocal leukoencephalopathy, paraneoplastic, Marchiafava-Bignami disease, Spongy degeneration, Alexander's disease, Canavan's disease, metachromatic leukodystrophy and Krabbe's disease. Ngr3 is useful for inducing an immune response in a mammal against Ngr3, as a bait protein in a two-hybrid or three-hybrid assay, and as a research tool for identifying, characterisation and purification of interacting, regulatory proteins. The nucleotide sequences of the invention are useful for screening for RFP associated with certain disorders, for genetic mapping, and for gene therapy. The vector containing Ngr3 is useful for producing non-human transgenic animals. The Ngr3 binding antibody is useful for isolating and purifying Ngr3, for localisation and/or quantitation of Ngr3, and for diagnostic and therapeutic purposes. The sequences of the invention, vectors and antibodies are useful for treating or preventing unregulated cellular growth such as cancer and tumour growth. This sequence represents the mature human Ngr3 protein of the invention.

Sequence 390 AA:

Query Match 93.5% Score 2082; DB 23; Length 390;
Best Local Similarity 100.0%; Pred. No. 1,4e-186;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

31 CPMCTCTSSPPYVSCQANFSSVPLSLPSTQRLFLQNNLITLRPGFGSLTLTLMF 90
1 CPMCTCTSSPPYVSCQANFSSVPLSLPSTQRLFLQNNLITLRPGFGSLTLTLMF 60
91 SNNLSTTPTTFPHLQALELDLGDNRHLRLSEPTFGCLERLSLHLYRCQLSLPGNI 150
61 SNNLSTTPTTFPHLQALELDLGDNRHLRLSEPTFGCLERLSLHLYRCQLSLPGNI 120
151 FRLVSLQLYLQENSLHLQDLDLADLANLISLFLHGRNRLTEHVRGSLDRLL 210
121 FRLVSLQLYLQENSLHLQDLDLADLANLISLFLHGRNRLTEHVRGSLDRLL 180
211 HGRNLOGVHRAARGLSRLTILYLFNNSLASLPGELADLPSEFLRLNANPWACDCAR 270
181 HGRNLOGVHRAARGLSRLTILYLFNNSLASLPGELADLPSEFLRLNANPWACDCAR 240
271 PLWAFORARVSSSDVTGATPPEGRGRDLRALREADFOACPPAATPRGSSRAKSSSNH 330
241 PLWAFORARVSSSDVTGATPPEGRGRDLRALREADFOACPPAATPRGSSRAKSSSNH 300
331 LYGVAAGAPPADPSTLYRDLPAEDSRGQGDAPTEDDYWGYGEGDQRGEMCPGAC 390
301 LYGVAAGAPPADPSTLYRDLPAEDSRGQGDAPTEDDYWGYGEGDQRGEMCPGAC 360
391 QAPDSRCPALSAQLPSPLCLLLLVPHNL 420
361 QAPDSRCPALSAQLPSPLCLLLLVPHNL 390

RESULT 3
ABG15173
ID ABG15173 standard: Protein; 807 AA.
XX
XX ABG15173;
AC
XX
XX 18-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #15164.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX
XX W0200175067-A2.

XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001MO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Drmanac RT, Liu C, Tang YT;
DR WPT: 2001-639362/73.
XX N-PSDB: AAS79360.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity.

Claim 20; SEQ ID NO 45532; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations in responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human CC diagnostic amino acid sequences. ABG00010-ABG30377 represent novel human CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 807 AA:

Query Match 88.8% Score 1977.5; DB 22; Length 807;
Best Local Similarity 91.9%; Pred. No. 2.4e-176;
Matches 385; Conservative 1; Mismatches 16; Indels 17; Gaps 4;

11 APASCLLMLLALPLAAPSCLMCTCYSSPPYVSCQANFSSVPLSLPSTQRLFLQNN 70
397 APASCLLMLLALPLAAPSCLMCTCYSSPPYVSCQANFSSVPLSLPSTQRLFLQNN 456
71 LITTLRPGTFGSLTLW-----LFSNNLSTTPTTFPHLQALELDLGDNRHLRLSL 122
457 LITTL-----GARLHWQPAHPVALLOQPLHL-PCHFPPQLALELDLGDNRHLRLSL 508
123 EPTFGELERLQSLHLRYRQSSLPFGNIFRGVLSQYLYLQENSLHLHQQDLFADLANLS 182
509 EPTFGELERLQSLHLRYRQSSLPFGNIFRGVLSQYLYLQENSLHLHQQDLFADLANLS 568
183 HL-FLHGRNRLRLTEHVRGSLSLDLRLHGRNLOGVHRAARFGLSRLTILYLFNNSLAS 241
569 PTFSLHGNRLRLTEHVRGSLSLDLRLHGRNLOGVHRAARFGLSRLTILYLFNNSLAS 628
242 LPEGLADLPSEFLRLNANPWACDCARPLWAFORARVSSSDVTGATPPEGRGRDLRA 301
629 LPEGLADLPSEFLRLNANPWACDCARPLWAFORARVSSSDVTGATPPEGRGRDLRA 688
302 LREADFOACPPAATPRGSSRAKSSSNHLYGVAAGAPPADPSTLYRDLPAEDSRGQ 361
689 LREADFOACPPAATPRGSSRAKSSSNHLYGVAAGAPPADPSTLYRDLPAEDSRGQ 748

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABC00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/publ/ishne_pc_sequences.

Sequence 241 AA;

Query Match	37.6%;	Score 836.5;	DB 22;	Length 241;
Best Local Similarity	91.0%;	Pred. No. 4.7e-70;		
Matches 161; Conservative	2;	Mismatches 7;	Indels 7;	Gaps 1;

172 DDLFADIANSHLEHGNRLTTEHVERGLGSJDRLLHGNRQGVHNAATROUSNLT 23

Db 1 DDLEADLANSHLFLHGNRLRLTEHVERGGLGLDRLLHGNRLQGVHKAARGLSKLI1

232 LYFNNSLASLPGALADLPSEFLRLNANPWACDCRRAPLWAMFORARVSSSDVICAP 291

61 LYFNNSLASLPGALADLPSEFLRLNANPWACDCRARPLWAWFQRRARVSSSDVTCATP 120

292 PERQGRDLREADFQACPPAAPTTRPCSRARGNSSSNHLYGVAEAGAPPADPSILI 340

Db 121 PERQGRDLALREGRFQACPPAGPTRPGSRARGNSFSNHLVG-----PSRGFIVY L/0

RESULT 7	
AAO21487	
ID	AAO21487 standard; Protein; 392 AA.

AC AA021487;

DT 15-AUG-2002 (first entry)

partial human NgrR3 protein sequence

Cerebroprotective; neuroprotective; cytoskeletal Nogo receptor; homologue
 KN; NGR; axonal growth; central nervous system; CNS; cerebral injury;
 KN; spinal cord injury; stroke; demyelinating disease; multiple sclerosis;
 KN; myofascial demyelination; encephalomyelitis; Marchiafava-Bignami disease
 KN; multifocal leukoencephalopathy; panencephalitis; Spong degeneration;
 KN; Alexander's disease; Canavan's disease; metachromatic leukodystrophy;
 KN; Krabbe's disease; immune; bait protein; genetic mapping; gene therapy;
 KN; transgenic animal; unregulated cellular growth; cancer; tumour; human

OS Homo sapiens

EH	Key	Location/Qualifiers
FT	Region	5..27
FT		/label= LRR1_region
FT	Region	28..51
FT		/label= LRR2_region
FT	Region	52..76
FT		/label= LRR3_region
FT	Region	77..100
FT		/label= LRR4_region
FT	Region	101..124
FT		/label= LRR5_region
FT	Region	125..148
FT		/label= LRR6_region

FT	Region	149..172
FT		/label= LRR7_region
FT	Region	173..196
FT		/label= LRR8_region
FT	Region	206..225
FT		/label= LRRCT_region
FT	Region	256..396
FT		/label= CTS_region
FT		/note= "CT Signalling region"
FT	Region	370..392
FT		/label= GPL_region

PF	06-OCT-2001; 2001WO-US31488.
XX	
PR	06-OCT-2000; 2000US-238361P.

PA (UYA) UNIV YALE.
PA (BIOT) BIOGEN INC.

PI Strittmatter SM, Cate RL, Sah DWY;

DR WPI; 2002-416677/44
DR N-PSDB; AAL38335.

Novel Nogo receptor homolog polypeptide, Ngr2 or Ngr3, useful for treating central nervous system disorder, cerebral injury, spinal cord injury, stroke, and demyelinating diseases -

PS Claim 17; Fig 3; 277pp; English.

XX The invention relates to a Nogo receptor homologue polypeptide, Ngr3 or NGR3, comprising a 50 amino acid LRRC1 sequence, a 284 amino acid NLRRC1 sequence, or a 420, 461 or 392 amino acid sequence, all given in the sequence, or a Ngr3 protein or its binding antibody is useful for specification. The Ngr3 protein or its binding antibody is useful for decreasing inhibition of axonal growth of a central nervous system (CNS) neuron, by contacting the neuron Ngr3 or its antibody, and for treating CNS disease, disorder or injury. Ngr3 or a vector comprising Ngr3 is useful for treating cerebral injury, spinal cord injury, stroke, demyelinating diseases, e.g. multiple sclerosis, monoplasic demyelination, encephalomyelitis, multifocal leukoencephalopathy, panencephalitis, Marchiafava-Bignami disease, metachromatic leukodystrophy and Alexander's disease. Ngr3 is useful for inducing an immune response in a Krabbe's disease. Ngr3 is a bait protein in a two-hybrid or three-hybrid mammal against Ngr3, as a bait protein for identification, characterisation and assay, and as a research tool for identification, regulatory proteins. The nucleotide purification of interacting, regulatory proteins. The nucleotide sequences of the invention are useful for screening for RFLP associated with certain disorders, for genetic mapping, and for gene therapy. The vector containing Ngr3 is useful for producing non-human transgenic animals. The Ngr3 binding antibody is useful for isolating and purifying Ngr3, for localisation and/or quantitation of Ngr3, and for diagnostic and therapeutic purposes. The sequences of the invention, vectors and antibodies are useful for treating or preventing unregulated cellular growth such as cancer and tumour growth. This sequence represents the partial human Ngr3 protein sequence of the invention.

SQ Sequence 392 AA

Query Match	36.4%	Score 811;	DB 23;	Length 392;
Best Local Similarity	51.2%	Pred No. 2, 3e-67;		
Matches 172; Conservative	40;	Mismatches 102;	Indels 22;	Gaps 5

58 LPPSTORLEQNNLIRTLPGTFCGNNLLTLMLFNNNLSSTIYPTGFRHQALEELDGDNR 117

Db 3 IPVDSEKVELQNNRIGLLQPGHFSPAMVTLWYSNNITYIHPSSTFEFGFVHLEELDGDNR 62

QY 118 HLRSEBDTFQGERLQSLLHYRCQSLSPONIFRGVLSQYLXLQENSLHLDQDLDFAD 170
 ||||| : : | :
 Db 63 QLRTPETFOGLVKHLALXYKCGISALPBGVFEGGHSLOLYXLQDNHIEYLQDDIFVD 122


```

Db      194 PERAFRLHSLDRLLLQNRYAVHVAHFAERDLGRIMTLVLFANNLSALPTEALAPLRAIQ 253
Oy      255 FLRLNANPMACDCRARPIMAFORARVSSDVATCPPEPQGRDLRLRADADQAC----- 310
        :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      254 YLRINDNPMWDCRCARPLMALKFRGSSSEVPCSLPQRLAKRLANDLQGCYVAT 313
Oy      311 -----PPAA-----PTRPGSRAR-----GNS 326
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      314 GPYHPITWGRATDEEPGLPKCCOPDAADKASYLEPGRPASAGNALGRVPPGDSPPGNC 373
Oy      327 SS-NHL-----YGVAEAGAPPADPSTLYR-----DLPAEDSRQG----- 361
        :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      374 GSPRINDNSPPTGLPGSAEP---PLTAVRPEGSEPPGFTSGPRRRPGCSRKNRTSRHCRL 431
Oy      362 GDAPTEDDYWGCGEGEDRGCEOMCPGACQAP 394
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      432 GQAGS-----GGGTGDESGALPILTCILTP 459

          1745
          AAY41745 standard: Protein: 473 AA.
          AAY41745:
          07-DEC-1999 (first entry)
          Human PRO526 protein sequence.
          Human PRO: EST; expressed sequence tag; PCR primer; hybridisation;
          KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
          secreted protein; transmembrane protein.
          Homo sapiens.
          OS
          PN W09946281-A2.
          PD 16-SEP-1999.
          PF 08-MAR-1999; 99MO-US05028.
          XX
          PR 10-MAR-1998; 98US-0077450.
          PR 11-MAR-1998; 98US-0077632.
          PR 11-MAR-1998; 98US-0077641.
          PR 12-MAR-1998; 98US-0077649.
          PR 13-MAR-1998; 98US-0077791.
          PR 17-MAR-1998; 98US-0078004.
          PR 20-MAR-1998; 98US-0040220.
          PR 20-MAR-1998; 98US-0078886.
          PR 20-MAR-1998; 98US-0078910.
          PR 20-MAR-1998; 98US-0078936.
          PR 20-MAR-1998; 98US-0078939.
          PR 25-MAR-1998; 98US-0079294.
          PR 26-MAR-1998; 98US-0079656.
          PR 27-MAR-1998; 98US-0079663.
          PR 27-MAR-1998; 98US-0079664.
          PR 27-MAR-1998; 98US-0079689.
          PR 27-MAR-1998; 98US-0079728.
          PR 27-MAR-1998; 98US-0079786.
          PR 27-MAR-1998; 98US-0079920.
          PR 30-MAR-1998; 98US-0079923.
          PR 30-MAR-1998; 98US-0080105.
          PR 31-MAR-1998; 98US-0080107.
          PR 31-MAR-1998; 98US-0080165.
          PR 31-MAR-1998; 98US-0080194.
          PR 31-MAR-1998; 98US-0080327.
          PR 01-APR-1998; 98US-0080328.
          PR 01-APR-1998; 98US-0080333.
          PR 01-APR-1998; 98US-0080334.
          PR 08-APR-1998; 98US-0081049.
          PR 08-APR-1998; 98US-0081070.
          PR 08-APR-1998; 98US-0081071.
          PR 09-APR-1998; 98US-0081195.

```

```

PR      09-APR-1998; 98US-0081203.
PR      09-APR-1998; 98US-0081229.
PR      15-APR-1998; 98US-0081817.
PR      15-APR-1998; 98US-0081838.
PR      15-APR-1998; 98US-0081952.
PR      15-APR-1998; 98US-0081955.
PR      21-APR-1998; 98US-0082568.
PR      21-APR-1998; 98US-0082569.
PR      22-APR-1998; 98US-0082700.
PR      22-APR-1998; 98US-0082704.
PR      22-APR-1998; 98US-0082804.
PR      23-APR-1998; 98US-0082767.
PR      23-APR-1998; 98US-0082796.
PR      27-APR-1998; 98US-0083336.
PR      28-APR-1998; 98US-0083322.
PR      28-APR-1998; 98US-0083392.
PR      29-APR-1998; 98US-0083495.
PR      29-APR-1998; 98US-0083496.
PR      29-APR-1998; 98US-0083499.
PR      29-APR-1998; 98US-0083500.
PR      29-APR-1998; 98US-0083545.
PR      29-APR-1998; 98US-0083554.
PR      29-APR-1998; 98US-0083558.
PR      29-APR-1998; 98US-0083559.
PR      30-APR-1998; 98US-0083742.
PR      05-MAY-1998; 98US-0084366.
PR      06-MAY-1998; 98US-0084414.
PR      06-MAY-1998; 98US-0084414.
PR      07-MAY-1998; 98US-0084498.
PR      07-MAY-1998; 98US-0084598.
PR      07-MAY-1998; 98US-0084600.
PR      07-MAY-1998; 98US-0084627.
PR      07-MAY-1998; 98US-0084637.
PR      07-MAY-1998; 98US-0084639.
PR      07-MAY-1998; 98US-0084640.
PR      07-MAY-1998; 98US-0084643.
PR      13-MAY-1998; 98US-0085323.
PR      13-MAY-1998; 98US-0085338.
PR      13-MAY-1998; 98US-0085339.
PR      15-MAY-1998; 98US-0085573.
PR      15-MAY-1998; 98US-0085579.
PR      15-MAY-1998; 98US-0085580.
PR      15-MAY-1998; 98US-0085582.
PR      15-MAY-1998; 98US-0085689.
PR      15-MAY-1998; 98US-0085697.
PR      15-MAY-1998; 98US-0085700.
PR      15-MAY-1998; 98US-0085704.
PR      18-MAY-1998; 98US-0086023.
PR      22-MAY-1998; 98US-0086392.
PR      22-MAY-1998; 98US-0086414.
PR      22-MAY-1998; 98US-0086430.
PR      22-MAY-1998; 98US-0086486.
PR      28-MAY-1998; 98US-0087098.
PR      28-MAY-1998; 98US-0087106.
PR      30-JUL-1998; 98US-0087208.
PR      11-SEP-1998; 98US-0100038.

PA      (GETH ) GENENTECH INC.
PI      Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
DR      WPI; 1999-551358/46.
DR      N-PSDB; AA234229.
XX      New secreted and transmembrane polypeptides and their polynucleotides,
PT      useful for treating blood coagulation disorders, cancers and cellular
PT      adhesion disorders -
PS      Claim 12, Fig 161; 530pp; English.
XX      The present invention describes secreted and transmembrane polypeptides
CC      and their polynucleotides. The nucleotide sequences are useful as
CC      sources of probes, primers, for chromosome mapping, and for generation

```


QY 327 SS-NHL-----YGAAGAGAPPADPSTLYR-----DLPAEDSRGRGQ----- 361
 Db 374 SGPRIINDSPFGLTPGSAEP--PLTAVRPEGSEPPGFTSGPRRRPGCSKRNTRSHCRL 431
 QY 362 GDAPTEDDYWGCGEDDQGEQMPGCAACQAP 394
 Db 432 GQAGS-----GGGCTGDSGSGALPSLTCSLTP 459

RESULT 11
 AAB24410
 ID AAB24410 standard; Protein: 473 AA.
 XX AAB24410;
 AC
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Human PRO526 protein sequence SEQ ID NO:127.
 XX
 XX Human: PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
 KM diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
 KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;
 XX cytosolic; gene therapy; vaccine.
 OS Homo sapiens.
 XX
 PN WO200032221-A2.
 XX
 XX 08-JUN-2000.
 PD
 XX 30-NOV-1999; 99WO-US28313.
 PF
 XX 01-DEC-1998; 98WO-US25108.
 PR 16-DEC-1998; 98US-0112850.
 PR 12-JAN-1999; 99US-0115554.
 PR 08-MAR-1999; 99WO-US05028.
 PR 12-MAR-1999; 99US-0123957.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-OCT-1999; 99US-0162506.

XX (GETH) GENENTECH INC.
 XX
 XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;
 PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
 PI Watanabe CK, Williams PM, Wood WI;
 XX
 XX WPI: 2000-412154/35.
 DR N-PSDB: AAA77613.
 XX
 PT Nucleic acids encoding PRO polypeptides useful for preventing,
 PT diagnosing and treating a cardiovascular, endothelial or
 PT angiogenic disorders in mammals -
 XX
 XX Claim 72; Fig 46; 315pp: English.
 XX
 XX The present invention describes nucleic acids encoding PRO polypeptides
 XX useful for preventing, diagnosing and treating a cardiovascular disorder
 XX in mammals by
 XX modulating cell proliferation, angiogenesis and cardiovascularisation,
 XX and for identifying agonists and antagonists of these processes. The
 XX nucleic acids and the proteins they encode may be used in the

CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate PRO expression such as cardiovascular, endothelial or
 CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
 CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
 CC containing them and the PRO polypeptide may be used to treat disorders
 CC associated with decreased PRO expression. AAA77510 to AAA77721 and
 CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
 CC the exemplification of the present invention.
 XX
 XX Sequence 473 AA:
 SQ
 Query Match 36.3%; Score 808; DB 21; Length 473;
 Best Local Similarity 42.2%; Pred. No. 5,7e-67;
 Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

QY 18 LMLALPLAPAGCPMCTCYSSP--TVSCQANNFSSVPLSLPSTORFLQNLIRTLR 76
 Db 15 VLWLQAMQVANP-CPGACVCYNEPKYTTSCPOQGLQAVPGVIAASQRIELHGNRISHVP 73
 QY 77 PGTEGS--NLITLMLFNNLSTYPTGFRHLQALBELDLDGNHRLSLDEPTGGLERLQ 134
 Db 74 AASFRACRNLTILMLHSNVLARIDAAFTGLLEDDLDSDNQLSVDPATFHGLGRHL 133
 QY 135 SLHLRQQLSPENIRGVSLQYLYLOENSLHLQDDLFADLANLSHLFHLGNRLRL 194
 Db 134 TLHDLRCGLQELPGRLAALQYLYLDNNAQALPDDYFRDLGNLTFLHGNRISV 193
 QY 195 TEHVEFGLSIDRLHLGNRLQGVHRAAFRLSLRTLLYLFNNSLASLPREALADPSE 254
 Db 194 PERAFRGLHSIDRLHLQNRVAHVHAFRLDGLRLMTLYLFANNLSALPREALAPRALQ 253
 QY 255 FLRLANPWCDCRARPIMANFQORAVSSDYTCATPEPQGRDLRALREADFOAC--- 310
 Db 254 YLRINDNPWCDCRARPIMAWLQKFRGSSSEPCSLPQRLAGRLKLANLDQCAVAT 313
 QY 311 -----PPAA-----PFRGSRAR-----GNS 326
 Db 314 GPHPIWTGRATDEEPLGLPKCCOPDADKASVLEPGRPASAGNALGRVPQDPPGNG 373
 QY 327 SS-NHL-----YGAAGAGAPPADPSTLYR-----DLPAEDSRGRGQ----- 361
 Db 374 SGPRIINDSPFGLTPGSAEP--PLTAVRPEGSEPPGFTSGPRRRPGCSKRNTRSHCRL 431
 QY 362 GDAPTEDDYWGCGEDDQGEQMPGCAACQAP 394
 Db 432 GQAGS-----GGGCTGDSGSGALPSLTCSLTP 459

RESULT 12
 AAY95345
 ID AAY95345 standard; Protein: 473 AA.
 XX
 XX AAY95345;
 AC
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE Human PRO526 antitumour protein.
 XX
 XX PRO526; human; antitumour; tumour; therapy; cytostatic;
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
 KW central nervous system cancer; melanoma; leukaemia; neoplasm.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..26
 FT Protein /label= Signal_peptide
 FT Protein /label= 27..473
 FT Domain /label= PRO526
 FT Modified-site 411..427
 FT /note= "von Willebrand factor type C domain"
 FT 82..86


```

OY 18 LLMALLPLAAPSCPMICTCYSSPP-TVSCANNSSVPLSPSTORLFLONNLRTIR 76
Db 15 VLMLOAMOVAAAP-CPGACVCYNEPKVTSCPOOGLAVPGIPAASQRIFLHGNRISHVP 73
OY 77 PGTGGS--NLTLWLFNSNLTSTIPGTFRIQLALDELQDNHRLSLRDPDFOGLERLQ 134
Db 74 AASFRACRNLTILMHSNVLARIDAAAFGLALDELQDLSDNAQLRSVDPATFHGGRH 133
OY 135 SLHLRYCOLSLSPGNIFRGVLSLOYLYLOENSLIHLQDDLFADLANLSLFLHGRNRL 194
Db 134 TLHLDRGLOELGPGFLRGVLAALQYLYLODNLQALPDPTFDDTLGNLTHLFLHGRNRSV 193
OY 195 TEHFVFGISLDRLLHGNRLQCVHRAAFRGSLRLLTYLFNNSLASLPGALADPLSL 254
Db 194 PERAFRGHLSLDRLLHGNRVAVHRAFRDGLRMTLYLFANNLSALPTALPLRALQ 253
OY 255 FLRLANPACDCRRARPLWAMPORARVSSDYTCATPREGRDRLRALREADFOAC---- 310
Db 254 YLRLNDNPMVCCRRARPLWAMLOKFRGSSSEVPCSLPQRLAGRDRLKRLAANDLOGCAVAT 313
OY 311 -----PPAA-----PTRGSRAR-----GNS 326
Db 314 GPYHPIWGRATDERPLPKCCOPDADKASVLEPRFASAGNALGRVPGDSPGNG 373
OY 327 SS-NHL-----YGVADAGAPPADPSTLYR-----DLPAEDSRGRG----- 361
Db 374 SGPRHINDSPGTLPGSAEP--PLTAVRPEGSEPPGTFSGRRRPGCSKRNTRSHCRL 431
OY 362 GDPAPDEDYMGVGGEGDGRGKMGCAACQAP 394
Db 432 GQAGS-----GGGCTGSESGALPSLTCSLTP 459

```

PR 18-FEB-2000: 2000MO-US04342.
 PR 22-FEB-2000: 2000MO-US04314.
 PR 24-FEB-2000: 2000MO-US04914.
 PR 15-MAR-2000: 2000MO-US06884.
 PR 20-MAR-2000: 2000MO-US07377.
 PR 21-MAR-2000: 2000MO-US07532.
 PR 30-MAR-2000: 2000MO-US08439.
 PR 17-MAY-2000: 2000MO-US13705.
 PR 22-MAY-2000: 2000MO-US14042.
 PA (GETH) GENENTECH INC.
 PI Ashkenazi AJ, Baker KP, Chan B, Goddard A, Godowski PJ, Gurney AL;
 PI Hebert C, Henzel W, Kabakoff RC, Shelton DL, Tumas D, Watanabe CK;
 PI Wood WI;
 DR WPI: 2001-025253/03.
 DR N-PSDB: MAC91467.
 PT Thity three nucleic acids encoding PRO polypeptides which are useful
 PT in the diagnosis and treatment of immune related disorders, e.g.
 PT systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 PT thyroiditis and diabetes mellitus -
 PS Claim 58; Fig 14; 218bp; English.
 CC The present sequence is one of thirty three novel PRO polypeptides.
 CC The PRO polypeptides, anti-PRO antibodies, agonists and
 CC antagonists are useful for treating and diagnosing immune related
 CC disorders such as systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis,
 CC systemic sclerosis, idiopathic chronic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune hemolytic
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central
 CC and peripheral nervous systems (such as multiple sclerosis, idiopathic
 CC demyelinating polyneuropathy or Guillain-Barre syndrome, and chronic
 CC inflammatory demyelinating polyneuropathy), hepatobiliary diseases
 CC (such as infectious, autoimmune chronic active hepatitis, and chronic
 CC biliary cirrhosis), autoimmune chronic active hepatitis, primary
 CC inflammatory bowel disease, gluten-sensitive enteropathy and Whipple's
 CC disease, autoimmune or immune-mediated skin diseases (such as bullous
 CC skin diseases, erythema multiforme, contact dermatitis, psoriasis),
 CC allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
 CC food hypersensitivity and urticaria), immunological diseases of the
 CC lung (such as eosinophilic pneumonitis, idiopathic pulmonary fibrosis
 CC and hypersensitivity pneumonitis), transplant associated diseases
 CC including graft rejection and graft-versus-host diseases.
 CC
 CC Sequence 473 AA:
 SQ

```

PN WO200073452-A2.
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000MO-US15264.
XX

```

```

PR 02-JUN-1999; 99MO-US12252.
PR 20-JUL-1999; 99US-0144732.
PR 20-JUL-1999; 99US-0144758.
PR 28-JUL-1999; 99US-0146232.
PR 01-SEP-1999; 99MO-US20111.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21347.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99MO-US28313.
PR 01-DEC-1999; 99MO-US28634.
PR 09-DEC-1999; 99MO-0170262.
PR 20-DEC-1999; 99MO-US30911.
PR 05-JAN-2000; 2000MO-US00319.
PR 06-JAN-2000; 2000MO-US00376.
PR 11-FEB-2000; 2000MO-US03565.
PR 18-FEB-2000; 2000MO-US04341.

```

Query Match 36.3%; Score 808; DB 22; Length 473;
 Best Local Similarity 42.2%; Pred. No. 5,706-67;
 Matches 191; Conservative 36; Mismatches 140; Indels 84; Gaps 12;
 OY 18 LLMALLPLAAPSCPMICTCYSSPP-TVSCANNSSVPLSPSTORLFLONNLRTIR 76
 Db 15 VLMLOAMOVAAAP-CPGACVCYNEPKVTSCPOOGLAVPGIPAASQRIFLHGNRISHVP 73
 OY 77 PGTGGS--NLTLWLFNSNLTSTIPGTFRIQLALDELQDNHRLSLRDPDFOGLERLQ 134
 Db 74 AASFRACRNLTILMHSNVLARIDAAAFGLALDELQDLSDNAQLRSVDPATFHGGRH 133
 OY 135 SLHLRYCOLSLSPGNIFRGVLSLOYLYLOENSLIHLQDDLFADLANLSLFLHGRNRL 194
 Db 134 TLHLDRGLOELGPGFLRGVLAALQYLYLODNLQALPDPTFDDTLGNLTHLFLHGRNRSV 193
 OY 195 TEHFVFGISLDRLLHGNRLQCVHRAAFRGSLRLLTYLFNNSLASLPGALADPLSL 254
 Db 194 PERAFRGHLSLDRLLHGNRVAVHRAFRDGLRMTLYLFANNLSALPTALPLRALQ 253
 OY 255 FLRLANPACDCRRARPLWAMPORARVSSDYTCATPREGRDRLRALREADFOAC---- 310
 Db 254 YLRLNDNPMVCCRRARPLWAMLOKFRGSSSEVPCSLPQRLAGRDRLKRLAANDLOGCAVAT 313

Fri Apr 4 12:29:25 2003

us-09-972-546-2.rag

```
Db 254 YLRINDNPWCDCRARPMLAMLOKFRGSSEVPCSLPORLAGRDLKRLANDIOGCVAAT 313
OY 311 -----PPAA-----PTRPGSRAR-----GNS 326
Db 314 GPYHPITWGRATDEEPLGLPKCCOPDADKASVLEPGRPASAGNALKGRVPPGDSPPGNG 373
OY 327 SS-NHL-----YGAEAAGAPPADPSTLYR-----DLPADSRGRG-----361
Db 374 SGPRIINDSPFGTLPGSAP--PLTAVRPEGSEPPGFPPTSGPRRRPGCSRKNRTRSHCRL 431
OY 362 GDAPTEDDYWG3YGGEDQGEOMCBGAACQAP 394
Db 432 GOAGS-----GGGTGTDSGSGALPILTCSLTP 459
```

Search completed: April 4, 2003, 12:09:55
Job time : 77 secs

APR 11 4, 2003, 12:07:27 ; Search time 35 Seconds

(average arguments)
2472.565 Million cell updates/sec

Title:	US-09-972-546-2
Perfect score:	2226

Sequence: 1 MLPGLRRLQAPASACLLM.....LSAGIPSPLLCTLLIVPHI. A30

Scoring table: BLOSUM62

671580 seqs, 206047115 residues

number of hits satisfying chosen parameters: 671580

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database

```

1:  sp.archaea:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp.human:*
5:  sp.invertebrate:*
6:  sp.mammal:*
7:  sp.mhc:*
8:  sp.organelle:*
9:  sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:
15: sp.rvirus:*
16: sp.bacterioph:*
17: sp.prtchcap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query No.	Score	Match	Length	DB	ID	Description
	1	815	36.7	473	6	Q9NOE3	Q9NOE3 macaca fasc
	2	808	36.3	473	4	Q9EZR6	Q9EZR6 homo sapien
	3	775	34.8	473	11	Q9PPI8	Q9PPI8 homo sapien
	4	762	34.2	473	11	Q9NM75	Q9NM75 mus musculu
	5	402.5	18.1	797	4	Q9UGS3	Q9UGS3 rattus norv
	6	384	17.3	811	4	Q75139	Q9UGS3 homo sapien
	7	374	16.8	481	4	Q9EQU5	Q75139 homo sapien
	8	369	16.6	485	4	Q9EQU5	Q9EQU5 homo sapien
	9	361.5	16.2	809	11	Q9DBH4	Q9DBH4 homo sapien
	10	360.5	16.2	578	11	Q8RM33	Q8RM33 mus musculu
	11	359.5	16.2	673	11	Q9CZT5	Q8RM33 rattus norv
	12	356.5	16.0	673	11	Q8RM25	Q9CZT5 mus musculu
	13	350	15.7	605	4	Q8TAY0	Q8TAY0 mus musculu
	14	343.5	15.4	567	11	Q9QZU3	Q9QZU3 homo sapien
	15	343	15.4	614	11	Q9D170	Q9QZU3 mus musculu
	16	341	15.3	562	11	Q8RI13	Q8RI13 mus musculu

17	341	570	11	Q9DBB9
18	337.5	15.2	640	4 Q9HCJ2
19	334.5	15.0	653	4 Q9HBM1
20	333.5	15.0	603	11 Q9DB17
21	331.5	14.9	430	4 Q13641
22	331	14.9	614	4 Q9EF55
23	331	14.9	614	6 Q9NM08
24	330	14.8	516	4 Q43300
25	329.5	14.8	687	11 Q9J110
26	326.5	14.7	603	11 Q70211
27	326.5	14.7	1480	5 Q9V7F8
28	326	14.6	799	5 Q9V964
29	325	14.6	426	11 Q9OYD9
30	321.5	14.4	649	4 Q9B685
31	320	14.4	1534	4 Q75093
32	319.5	14.4	648	11 Q9V583
33	319.5	14.4	737	5 Q9VU11
34	317.5	14.3	426	11 Q9Z0L0
35	317	14.2	1531	11 Q9VWB5
36	315.5	14.2	615	5 Q961X3
37	315	14.2	788	4 Q8W422
38	312.5	14.0	733	5 Q24250
39	312.5	14.0	1504	5 Q9XVY4
40	312.5	14.0	1504	5 Q9V7F9
41	311.5	14.0	1530	11 Q9NWG5
42	310	13.9	581	4 Q8T656
43	310	13.9	1531	11 Q8E279
44	309.5	13.9	544	11 Q8B723
45	308.5	13.9	572	4 Q960H3

ALIGNMENTS

RESULT

ID	Q9N0B3	PRELIMINARY:	PRI:	473 AA.
DT	01-OCT-2000 (TREMBlrel. 15, Created)			
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)			
OS	Unnamed protein product.			
OS	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
OC	Cercopithecinae; Macaca.			
NCBI_TaxID	9541;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirao M., Terao K.			
RA	Suzuki Y., Sugano S., Hashimoto K.:			
RT	"Isolation of full-length cDNA clones from macaque brain cDNA			
RL	libraries.";			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB045987; BAB01569.1;			
DR	InterPro: IPR001611; LRR.			
DR	InterPro: IPR000483; LRR_Cterm.			
DR	InterPro: IPR000372; LRR_Nterm.			
DR	InterPro: IPR003592; LRR_out.			
DR	InterPro: IPR003591; LRR_typ.			
DR	Pfam: PF00560; LRR; 8.			
DR	Pfam: PF01463; LRRCT; 1.			
DR	SMART; SM00370; LRR; 4.			
DR	SMART; SM00082; LRRCT; 1.			
DR	SMART; SM00013; LRRNT; 1.			
DR	SMART; SM00369; LRR_typ; 2.			
SO	SEQUENCE	473 AA; 5064 MW; 5329DDE83BD12C83 CPGC66.		

Query March	36.7%:	Score 816:	DB 6:	Length 473:
Best Local Similarity	42.6%:	Pred. No. 4	1e-61	
Matches 193: Conservative	36:	Mismatches	140	

QY 18 LMLALPLAPSCMLCTCYSSP-TVSCANNFSSVPLSLPPTQRLFLQNNLIRTLR 76

[illegible]

Db	1	MEARALRLLLVOCCLAPLEAPLAPCPREKDC-QIRPHULLCTNRGLRVAVPTSSLPSPH	59
Oy	61	-----STORLEFLQNNLIRTPRPFTGC--SNLLIUMLESNLL	94
Db	60	DVLVLSYSGNFTITTAFAFDHRIQLRGLDLOYNQIRSLHPKTFEKLRLSEELYSLENNLL	119
Oy	95	STIYGFGRHLQALEELDLDDNRHLRLSEDPDTQGERLQSLHLYRCQSLSPGNIFRGL	154
Db	120	QALAPGLAPLRLKRLILYANGN-EISRLSNGSEFGLSELYKLRLDGNALCALPDVAFAPL	178
Oy	155	VSLOLYLQE-----	164
Db	179	GNLIYLHESNRIRIFLGKNAFAQLKRLFLNLSANELQPSLRHAATFAPRLSLSSLSA	238
Oy	165	NSLIHLQDLPADLANLSHLEFLHNRRLLTTEHYFRGLSIDRLLLHGNR-----	214
Db	239	NSLOHLGRIFRIFHLPRGLSLSRNQTLTHLAPEAFNCEALRELRLEGNLSQLPALTE	298
Oy	215	-----LOGVHRAAFGLSLRITLLYLFNNSLASLPGEALADLPSEFLRLNA	260
Db	299	PLHSLDALDSGNELSAHPPTFGLHGRRLSLRRNALSLASGDDIRPAASPALYRLDLG	358
Oy	261	NEWACDCRAPL--WA--WFORAVSSSDYVCAIPRPENOGDLEALREAPDQACPPAPT	316
Db	359	NGMTDCDRIGRKPRMGDMHSQGRLLTYVQCRHPPLRLKGLTYLDLDDQLQNSCADPS	418
Oy	317	RPGSRARGSSNNHLYGVAEGAPADPSTYLRDLPAEDSRNGOG---GDAPTEDDYGC	372
Db	419	-PSASITADRRQPL--PTAAGEEMTPAGLAELPLPQPOLQOGRFLAGVA-----WD	469
Oy	373	GYGE-----DQRC---EQMCPACQACPP-----DSRG-----PALSA	403
Db	470	GARELVGNRSALRLSRROGLQDPSVAAAGAPQSLDLHKPKQGRPTRADPALAE	529
Oy	404	GLPS 407	
Db	530	PPTP 533	
RESULT 7			
Oy	09GZU5	PRELIMINARY: PRT: 481 AA.	
ID	09GZU5		
AC	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	leucine-rich repeat protein (Nyc1alopin).		
GN	CLRP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RA	PubMed=11062472;		
RA	Pusch C.M., Zeltz C., Brandau O., Pesch K., Achatz H., Feil S.,		
RA	Scharfe C., Maurer J., Jacobi F.K., Plinckers A., Andraasson S.,		
RA	Hardcastle A., Wissinger B., Berger W., Meiseld A.;		
RT	"The complete form of X-linked congenital stationary night blindness		
RT	is caused by mutations in a gene encoding a leucine-rich repeat		
RL	protein.";		
RL	Nat. Genet. 26:324-327(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=20517340; PubMed=11062471;		
RA	Bech-Hansen N.T., Naylor M.J., Maybaum T.A., Sparkes R.L., Koop B.,		
RA	Birch D.G., Bergen A.A., Pijnsen C.F., Polomeno R.C., Gal A.,		
RA	Drack A.V., Murell M.A., Jacobson S.G., Young R.S., Weleber R.G., et al.		
RT	"Mutations in NYX, encoding the leucine-rich proteoglycan nyc1alopin		
RT	cause X-linked complete congenital stationary night blindness.";		
RL	Nat. Genet. 26:319-323(2000).		
DR	EMBL; AJ278865; CAC19014.1; -.		
DR	EMBL; AF234868; AAC42685.1; -.		

RP SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=LUNG;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner U., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustingstein S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Rodriguez I., Sakanoto N.,
 RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakanoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker G., Wilmink L.,
 RA Wynchaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 EMBL: AK004681; BAB23469.1;
 MGD: MGI:1914123; 1200009022Rik.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00560; LRR; 10.
 DR Pfam: PF01463; LRRCT; 1.
 DR PRINTS: PRO0019; LEURICHRPT.
 DR SMART: SM00370; LRR; 8.
 DR SMART: SM00082; LRRCT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_TTP; 11.
 DR SEQUENCE 809 AA; 88809 MW; 5E86F5B8AE419FA CRC64;
 Query Match 16.2%; Score 361.5; DB 11; Length 809;
 Best Local Similarity 28.9%; Pred. No. 3,7e-22;
 Matches 146; Conservative 49; Mismatches 174; Indels 137; Gaps 19;
 QY 22 LALPLAAPS-CPMLCTCYSSPTYSQCOANPSSVP-LSLPPSTP----- 63
 DB 17 LAFPPRAESVCPERCDC-QHPQHLCTNRGLRAVPKTSLSLPDQVLYTSLGCGFTINIF 75
 QY 64 -----RLFLQNNLIRTLRPGTFG--SNLITLMLSSNNISTYPTGTFRLQALAE 110
 DB 76 AEDFRLGQLRLDQVQIRSLHPKTFEKLRLLELYLGNLILQALVPGTLARLRLRI 135
 QY 111 LDLDGNRLRLSLPTFGGLERLQSLHLRYQQLSLPLGNIIFRGVLSQVLYLQE----- 164
 DB 136 IYANGN-EIGRLSRGSGFGLSLVLYKRLDGNVGLPDAVPAPLGNLILYHLEENRIRFL 194
 QY 165 -----NSLIHLODDLPADLAN 180
 DB 195 GNAASQGLKRLFLMLSANLQPSLRHATFVPLRSITLLSANSLQHLGPRVFOHLPR 254
 QY 181 LSHFLHGNRLRLTEHVFRGLGSLDRLLHGNR-----LQ 216
 DB 255 LGLSLISGQNLTHLAPBEAFWGLEALRELRLEGNRLNQLPLTLLEPLHSLDALDSGNLS 314
 QY 217 GVHRAFRRLSLTLTYLFNNLSIASPGFALADLPSEFLRLNANPMACDGRAPL--WA 274
 DB 315 ALHPATFGHGRRLRELISLDMNLSALSGDIFAPALVRLDLDGGMWCDRLGLKRFMM 374
 QY 275 --WFORARYSSDYTCATPPEERGDRALREADPQ--AC--PPAATPRPGSRAGNSS 328
 DB 375 GNNHSGRLITLVYQCRRHPALRGKLYDLDDQLQNGSCVDPSPSP--AGSRQWPLPFL 433
 QY 329 NHIYGAEGAGAPPADPSTLYRLDPAE-----DSRGR-QGDAPTEDDYGGYGGEDORGE 382

DB 434 -----SEEGMTP--PAGLSQELPIQPOPOQRGRLLPVNA-----WGG-AAKELVGN 478
 QY 383 QMCPGAACQAPPDSRCPALISAGLPSP 408
 DB 479 RSLRLSRGPGPHGQPSAAAGSAP 504
 RESULT 10
 ID 08R5M3 PRELIMINARY; PRT: 578 AA.
 AC 08R5M3;
 DT 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE lib.
 GN Rattus norvegicus (Rat).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR RATS; TISSUE=BRAIN;
 RX MEDLINE=21645900; PubMed=11785964;
 RA Satoh K., Hata M., Yokota H.;
 RT "A Novel Member of the Leucine-Rich Repeat Superfamily Induced in Rat
 Astrocytes by beta-Amyloid."
 RL Biochem. Biophys. Res. Commun. 290:756-762(2002).
 EMBL: AB071036; BAB84586.1;
 DR SEQUENCE 578 AA; 64127 MW; 991BD057F5912591 CRC64;
 Query Match 16.2%; Score 360.5; DB 11; Length 578;
 Best Local Similarity 33.9%; Pred. No. 2.9e-22;
 Matches 106; Conservative 40; Mismatches 132; Indels 35; Gaps 9;
 QY 49 NNFSSVPL---SLPSTQRLFLQNNLIRTLRPGTF--GSNLTLMLSNNISTYPTGTF 103
 DB 207 NKLSDIPMGTFDALGNLQELQENQITLSPGLFHNHNRRLQRLYLSNNHISQLPQIFM 266
 QY 104 HLOALEEDLDGNRLRLSLPTFGGLERLQSLHLRYQQLSLPLGNIIFRGVLSQVLYLQ 163
 DB 267 QLPQNKLTLLFGN-SLRELSPGVFGPMPLRLMLYNNHTISLADNFTSHLNQVLI 325
 QY 164 ENSLIHLODDLPADLANSHLFLHGNRLRLTEHVFRGLGSLDRLLHGNRLQGVHRAAF 223
 DB 326 HNOQLYTISPGANGLTNLRELISLHNNALODDSDNFRSLANQNTSIQSNRLRLQPGSIF 385
 QY 224 RGLSRLTLTYLFNNLSIASPGFALADLPSEFLRLNANPMACDGRAPL--WAMFORARY 281
 DB 386 ANVNGLTITQLONNLENPLIGIFPHLVNLCRLYDNPWRCDSDILPHNNLLNRLARL 445
 QY 282 SSSDY-TCATPPEERGDRALREADPQAPPAATPRPGSRAGNSSNHIYGAEGAP 340
 DB 446 GTDTLPVCSPPANVRQSL-VIININE-----PPSVQGET----- 481
 QY 341 PADPSTLYRDLPA 353
 DB 482 PEYPS--YPPTPS 492
 RESULT 11
 ID 09CZT5 PRELIMINARY; PRT: 673 AA.
 AC 09CZT5;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE 2610528G05Rik protein.
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Akawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Flischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl L.M., Steadil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Schmitt L.M., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Kono H., Balderelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guenrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012169; BAB28075.1; -.
DR HSSP; P00740; 1EDM.
DR MGD; MGI:1921457; 2610528G05Rik.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR003961; FN-III.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_tyr.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00041; fn3_1.
DR Pfam; PF00360; LRR_6.
DR Pfam; PF01463; LRRT; 1.
DR Pfam; PF01462; LRRT; 1.
DR PRINTS; PRO0019; LEURICHRPT.
DR SMART; SM00181; EGF_1.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00060; FN3_1.
DR SMART; SM00370; LRR_3.
DR SMART; SM00082; LRRT; 1.
DR SMART; SM00013; LRRT; 1.
DR SMART; SM00369; LRR_tyr; 7.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR SEQUENCE 673 AA; 72336 MW; 9C53P90ADF43FBD9 CRC64;

Query Match 16.2%; Score 359.5; DB 11; Length 673;
Best Local Similarity 29.9%; Pred. No. 4.4e-22;
Matches 141; Conservative 40; Mismatches 201; Indels 89; Gaps 15;

OY 17 LLLMLALPLAAPSCPMCTCYSSPTVSCQANNFSSVPLSLPSTQRLFNQNLIRTLR 76
DB 11 LLLLVLLGSGVGGCPSCQC-NPQTVCTARQGTVPDPVDTGLYIFENGITITLD 69
OY 77 PGTFGS-----NLTLMLFSSNNLSTYPTGFRHLQALE 110
DB 70 VGCFFAGLPGQLDLSQNGITSLPGCIFQPLVNLSDLTANKLHETISNETFGRRLER 129
OY 111 LDGDNHRLSLRPTFOGLER-----LQSLHYRCQLSSLPGNITRGL--- 154
DB 130 LYLGNKR-IRHTQGFADALDRLELKPNDNELRVLPRLHPRLLDLSHNSIPALEAG 188
OY 155 -----VSLQYLYQENSLLHODDLFADLANLSHLFGNRLRLTTH---VERGLGSLDR 207
DB 169 ILDTANVEALRLAGLRLQDDEGLFGRLLNLHLDVYDQNL-----EHMPVYQGLRGLTR 244
OY 208 LLLHGN-RLOGVHRAAFGLSLRITLYLFNNSLASLPGEALADLPSEFLRLANPNACD 266
DB 364 STIREPTLSTSSQAPLTPPTTQASTVLTARPTMRAPODDCFASICLNGSCRL 423

OY 245 LRLAGNTRIAQIRPEDLAGITLALQELDVSNLSLQALPSDLSLFPRLRLAARNPNCIL 304
OY 267 CRARPLMAMFORARV---SSSDVTCATPPEROGRODLRALREADFOACRPAATP-----R 317
DB 305 CPLSMFGPMVRENHVVVLASPEETRGHFPPKKNAGRLLLDIDYADF-GCPVTTTATVPTR 363
OY 318 PGRARGNSSNNHLYGVAEAGAPRADPSTLYRDLPAEDSRGQGDAPTEDDYMGYGC-- 375
DB 364 STIREPTLSTSSQAPLTPPTTQASTVLTARPTMRAPODDCFASICLNGSCRL 423
OY 376 GEDORGEOMCP-----GAACQAP-----PDSRGRLSGLRPLCL 412
DB 424 GARNHNECSPRGFTGLYCESVEYDQMKPSSITPDT-----PRDPLRL 467

RESULT 12
ID Q8R2G5 PRELIMINARY; PRT; 673 AA.
AC Q8R2G5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE S11c-like 2 protein precursor.
GN S11c2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Schrewe H., Kutejova E.;
RT "Structure and expression analysis of the mouse S11c-like 2 (S11c12)
RT gene.";
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 673 S11c-like 2 PROTEIN.
FT SEQUENCE 673 AA; 72261 MW; AABDBA82DA8E9D32 CRC64;

Query Match 16.0%; Score 356.5; DB 11; Length 673;
Best Local Similarity 29.9%; Pred. No. 7.8e-22;
Matches 141; Conservative 39; Mismatches 202; Indels 89; Gaps 15;

OY 17 LLLMLALPLAAPSCPMCTCYSSPTVSCQANNFSSVPLSLPSTQRLFNQNLIRTLR 76
DB 11 LLLLVLLGSGVGGCPSCQC-NPQTVCTARQGTVPDPVDTGLYIFENGITITLD 69
OY 77 PGTFGS-----NLTLMLFSSNNLSTYPTGFRHLQALE 110
DB 70 VGCFFAGLPGQLDLSQNGITSLPGCIFQPLVNLSDLTANKLHETISNETFGRRLER 129
OY 111 LDGDNHRLSLRPTFOGLER-----LQSLHYRCQLSSLPGNITRGL--- 154
DB 130 LYLGNKR-IRHTQGFADALDRLELKPNDNELRVLPRLHPRLLDLSHNSIPALEAG 188
OY 155 -----VSLQYLYQENSLLHODDLFADLANLSHLFGNRLRLTTH---VERGLGSLDR 207
DB 169 ILDTANVEALRLAGLRLQDDEGLFGRLLNLHLDVYDQNL-----EHMPVYQGLRGLTR 244
OY 208 LLLHGN-RLOGVHRAAFGLSLRITLYLFNNSLASLPGEALADLPSEFLRLANPNACD 266
DB 245 LRLAGNTRIAQIRPEDLAGITLALQELDVSNLSLQALPSDLSLFPRLRLAARNPNCIL 304
OY 267 CRARPLMAMFORARV---SSSDVTCATPPEROGRODLRALREADFOACRPAATP-----R 317
DB 305 CPLSMFGPMVRENHVVVLASPEETRGHFPPKKNAGRLLLDIDYADF-GCPVTTTATVPTR 363
OY 318 PGRARGNSSNNHLYGVAEAGAPRADPSTLYRDLPAEDSRGQGDAPTEDDYMGYGC-- 375
DB 364 STIREPTLSTSSQAPLTPPTTQASTVLTARPTMRAPODDCFASICLNGSCRL 423

```


RC STRAIN-C57BL/6J: TISSUE-TESTIS;
RX MEDLINE-21085660; PubMed-11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Aachji J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Balasov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schliml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Noridone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
"Functional annotation of a full-length mouse cDNA collection.";
Nature 409:685-690(2001).
DR EMBL: AK027262; BAB32403.1; -;
DR HSSP: P23945; 1XUN.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF00560; LRR; 9.
DR Pfam: PF01463; LRRCT; 1.
DR Pfam: PF01462; LRRNT; 1.
DR SMART: SM00409; IG_1.
DR SMART: SM00408; IG2; 1.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 9.
KW Immunoglobulin domain.
SQ SEQUENCE 614 AA; 6910 MW; 41CFE40C21335681 CRC64;

Query Match 15.4%; Score 343; DB 11; Length 614;
Best Local Similarity 28.5%; Pred. No. 9,8e-21;
Matches 115; Conservative 50; Mismatches 133; Indels 106; Gaps 11;
1 MLPGRLRLLOAPASAC---LLMLLALPL--AAPSCPMCTCYSSPPVSCQANNFSSVP 55
1 MLAGGRMSHPSPLLACWOPILLVLGSLVSGSATCCPRCECSAQDRAVLCRRKRFVAVP 60
56 LSLPSTORLFLQNNLIRTL-----RPGTFGS--NLTLWL 89
61 EGIPETRLDGLGKNIKITLNDDEFASFPHELELENENIYSAVEPGAFNNLFLNLTGL 120
90 FSNNLSTIYPTGFRHQLAELELDGDN-----HLRSLEPD-----T 126
121 RSNRLKLIPLGVTGLSNLTGLDISENKIVILLDYMFOQLYNLKSLEVGNDLVYISHRA 180
127 FQGLERLOSILHYRCQLSSLPQINIFRGIVSLQYLYLOENSLHLQDDLFADLANLSHFL 186
181 FGLNSLEGLTLEKCNLTSLPEALSHLGLIVLRLRLNINAIIRDYSFKRLYRLKYLEI 240
187 -HGNRLRLTEHVFRL-----GSLDRLL 209
241 SHWPYLDITMTPNCLTGLMLTSLTSHCNLTAVPYLAVRHLVYLRFLNLSYNPITIGISM 300
210 LH-----GNRLGVHRAAFRGSLRTTIYLFNNSLASLPGALADLPSPLEFLRL 258
301 LHELLRLQELIOLVGCQLAVVEPYAFRGILNLYLRVLNVSGNQLTTLSESAFHSVGNLTFLIL 360

OY 259 NANPWACDCRARPPLWAMFORARV--SSDYTCATPPPERGDR 300
DB 361 DSNPLACDCRL--LWVFRRRRLNFRNQPTCATPPEFVGCKEKK 402

Search completed: April 4, 2003, 12:08:15
Job time : 40 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 10.3704 Seconds
(without alignments)
199.975 Million cell updates/sec

Title: US-09-972-546-2_COPY_261_310
Perfect score: 284
Sequence: 1 NPVACDCRRAPLWAFQRRAR.....PPERGRDLRALREADFCAC 50

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
1 number of hits satisfying chosen parameters: 10105

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwisssProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	18.7	100	1 MERT_STRLI	P30345 streptomyc
2	47	16.5	91	1 MERT_ALCSP	P94186 alcaigenes
3	46	16.2	91	1 MERT_SALTI	P94701 salmonella
4	45	15.8	98	1 ACYO_CHICK	P07032 gallus gall
5	44	15.5	85	1 YKK2_CAEEL	P34279 caenorhabdi
6	42.5	15.0	100	1 POL_HV1J3	P12498 human immun
7	42	14.8	91	1 MERT_PSEAF	P04131 pseudomonas
8	42	14.8	91	1 MERT_PSEAF	P04131 pseudomonas
9	42	14.8	91	1 MERT_PSEAF	P04131 pseudomonas
10	42	14.8	91	1 MERT_PSEAF	P04131 pseudomonas
11	42	14.8	91	1 MERT_PSEAF	P04131 pseudomonas
12	41.5	14.6	100	1 VP52_BPAPS	O26430 manduca sex
13	41	14.4	71	1 KEB2_ECOLI	O91196 bacterioph
14	41	14.4	88	1 V630_BPNM2	P13965 escherichia
15	40	14.1	90	1 VNS1_TACAO	O64224 mycobacteri
16	40	14.1	90	1 VNS1_TACAO	P26148 influenza a
17	39.5	13.9	72	1 VHEB_BPPL	P40639 alligator m
18	39.5	13.9	94	1 AESA_ALAMI	O07481 bacterioph
19	39	13.7	90	1 Y172_BURCE	P32578 burkholderi
20	39	13.7	98	1 VES3_BPV4	P12461 notophthal
21	38.5	13.6	65	1 VYIS_BPPH8	P08346 bovine pap
22	38.5	13.6	92	1 NEU2_HORSE	P05998 bacterioph
23	38	13.4	74	1 HMAA_AEDAE	P29552 aedes cabal
24	38	13.4	81	1 YLOC_BACSU	O31738 bacillus su
25	38	13.4	86	1 YLVM_ECOLI	P13048 escherichia
26	38	13.4	87	1 TWCE_BACSU	P39603 bacillus su
27	38	13.4	94	1 CH10_LISTN	O92949 listeria in
28	38	13.4	94	1 CH10_LISTN	O92949 listeria in
29	37.5	13.2	51	1 MT2_CANGA	O94611 listeria mo
30	37.5	13.2	72	1 NOD1_PHLIT	P41143 candida gla
31	37.5	13.2	87	1 DB15_BOVIN	O6m233 bos taurus
32	37	13.0	52	1 SRY_AKONZ	P36388 akodon azar
33	37	13.0	80	1 RRI6_LOTJA	P58125 lotus japon

34	37	13.0	85	1 RRI6_ORYSA	P12151 oryza sativ
35	37	13.0	88	1 RRI6_SINAL	P10359 sinapis alb
36	37	13.0	100	1 BOP1_BOVIN	P83107 bos taurus
37	37	13.0	100	1 BXA2_SAMCY	P33717 samia cynth
38	36.5	12.9	84	1 GLRX_PSEAF	O9nu55 pseudomonas
39	36.5	12.9	87	1 DB15_RAT	P56702 ratius robu
40	36	12.7	42	1 TXRO_ATPRO	P01478 atrox robu
41	36	12.7	64	1 NXS1_BUNFA	P10808 bungarus fa
42	36	12.7	67	1 YPE1_NPVLD	P36866 lymantria d
43	36	12.7	74	1 HMAA_APIME	P15856 epis mellif
44	36	12.7	82	1 Y696_BORBU	O51639 borrelia bu
45	36	12.7	88	1 YPE4_NPVLD	P36868 lymantria d

ALIGNMENTS

RESULT 1
ID MERT_STRLI STANDARD: PRT: 100 AA.

AC P30345: 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Mercuric transport protein (Mercury ion transport protein).
OS MERT.
GN Streptomyces lividans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.

OX NCBI_TaxId:1916;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:66 / 1326;
RX MEDLINE:93156687; PubMed:1494353;
RA Sedimeter R., Altenbuchner J.;
RT "Cloning and DNA sequence analysis of the mercury resistance genes of Streptomyces lividans."
RI Mol. Gen. Genet. 236:76-85(1992).

RL -1- FUNCTION: INVOLVED IN MERCURIC TRANSPORT. PASSES A HG(2+) ION FROM THE PERIPLASMIC MERT PROTEIN TO THE MERCURIC REDUCTASE (MERT).
CC FROM THE PERIPLASMIC MERT PROTEIN TO THE MERCURIC REDUCTASE (MERT).

-1- SUBCELLULAR LOCATION: Integral membrane protein.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----
DR EMBL: X65467; CAA46463.1; -
DR PIR: S23611; S23611.
DR PIR: S30171; S30171.
KW Transport; Mercury; Mercuric resistance; Transmembrane.

FT TRANSMEM 19 39
FT TRANSMEM 19 39
FT METAL 43 63
FT METAL 31 31
FT METAL 32 32
FT METAL 82 82
FT METAL 83 83
SQ SEQUENCE 100 AA: 10353 MW: 55654.00DA4CC215E4 CRC64;
Best Local Similarity 29.2%; Score 53; DB 1; Length 100;
Matches 14; Conservative 4; Mismatches 26; Indels 4; Gaps 1;

Qy 1 NPV-----ACCCRRAPLWAFQRRARVSSDYTCATPPERGRDLRALRE 44
Db 52 SPWLLRAVAVLLAGALTWLRRRRRTGNCACSLPARPTQNDHDLRK 99
RESULT 2

MERP_ALCSP STANDARD; PRT: 91 AA.
 ID MERP_ALCSP
 AC P94186;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mercuric transport protein periplasmic component precursor
 DE (periplasmic mercury ion binding protein) (Mercury scavenger protein).
 GN MERP.
 OS Alcaligenes sp.
 OG Plasmid IncH2 PMER610.
 CC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 CC Alcaligenes.
 CC NCBI_TaxID=512;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97303088; PubMed=9159519;
 RA Yurleva O., Kholodil G., Minakhin L., Gorlenko Z., Kalyaeva E.,
 RA "Intercontinental spread of promiscuous mercury-resistance
 RA transposons in environmental bacteria."
 RA Mol. Microbiol. 24:321-329(1997).
 CC -1- FUNCTION: MERCURY SCAVENGER THAT SPECIFICALLY BINDS TO ONE HG(2+)
 CC ION AND WHICH PASSES IT TO THE MERCURIC REDUCTASE (MERA) VIA THE
 CC MERT PROTEIN.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (Probable).
 CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y08993; CAA70197.1; -
 DR HSSP: P04129; IAFJ.
 DR InterPro: IPR001802; HG_scsavenger.
 DR InterPro: IPR001934; HeavyMe_transpt.
 DR Pfam: PF00403; HMA_1.
 DR PRINTS: PRO0946; HGSCAVENGER.
 DR PROSITE: PS01047; HMA_1; 1.
 DR PROSITE: PS00846; HMA_2; 1.
 DR Transport; Mercuric resistance; Periplasmic; Metal-binding; Signal;
 KW Plasmid.
 CC
 CC SIGNAL 1 19 POTENTIAL.
 CC CHAIN 20 91 MERCURIC TRANSPORT PROTEIN PERIPLASMIC
 CC COMPONENT.
 FT DOMAIN 23 89 HMA.
 FT METAL 33 33 HG(2+) (POTENTIAL).
 FT METAL 36 36 HG(2+) (POTENTIAL).
 SQ SEQUENCE 91 AA: 9505 MW: 93988DE054D8AA7F CRC64:
 Query Match 16.5%; Score 47; DB 1; Length 91;
 Best Local Similarity 40.9%; Pred. No. 24;
 Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 QY 11 PLWAMFORARVSSDYCATPP 32
 Db 16 PVMATQTVTLSPQMTCTSP 37
 RESULT 3
 ID MERP_SALTI STANDARD; PRT: 91 AA.
 AC P94701; 008125;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mercuric transport protein periplasmic component precursor
 DE (Periplasmic mercury ion binding protein) (Mercury scavenger protein).

GN MERP OR HCM1.153.
 OS Salmonella typhi.
 OS Enterobacter agglomerans (Pantoea agglomerans), and
 OS Enterobacter cloacae.
 OG Plasmid PHCM1, Plasmid PKLH272, and Plasmid PKLH256.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 CC NCBI_TaxID=601, 549, 550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S. typhi; STRAIN-CT18; PLASMID-PHCM1;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dow L., White N., Farrar J.,
 RA Feltham T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Kirogh A., Larsen T.S., Leather S., Mole S., O'Garra P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RA "Complete genome sequence of a multiple drug resistant *Salmonella*
 RA enterica serovar Typhi CT18";
 RT Nature 413:848-852(2001).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP SPECIES-E. agglomerans, and E. cloacae; STRAIN-KH72, and TC256;
 RC PLASMID-PKLH272, and PKLH256; TRANSPOSON-Tn5036;
 RX MEDLINE=97303088; PubMed=9159519;
 RA Yurleva O., Kholodil G., Minakhin L., Gorlenko Z., Kalyaeva E.,
 RA "Intercontinental spread of promiscuous mercury-resistance
 RA transposons in environmental bacteria."
 RA Mol. Microbiol. 24:321-329(1997)
 CC -1- FUNCTION: MERCURY SCAVENGER THAT SPECIFICALLY BINDS TO ONE HG(2+)
 CC ION AND WHICH PASSES IT TO THE MERCURIC REDUCTASE (MERA) VIA THE
 CC MERT PROTEIN.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (Probable).
 CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL51383; CAD09748.1; -
 DR EMBL: Y08992; CAA70182.1; -
 DR EMBL: Y09025; CAA70229.1; -
 DR HSSP: P04129; IAFJ.
 DR InterPro: IPR001802; HG_scsavenger.
 DR InterPro: IPR001934; HeavyMe_transpt.
 DR Pfam: PF00403; HMA_1.
 DR PRINTS: PRO0946; HGSCAVENGER.
 DR PROSITE: PS01047; HMA_1; 1.
 DR PROSITE: PS00846; HMA_2; 1.
 DR Transport; Mercuric resistance; Periplasmic; Metal-binding; Signal;
 KW Plasmid; Complete proteome.
 CC
 CC SIGNAL 1 19 POTENTIAL.
 CC CHAIN 20 91 MERCURIC TRANSPORT PROTEIN PERIPLASMIC
 CC COMPONENT.
 FT DOMAIN 23 89 HMA.
 FT METAL 33 33 HG(2+) (POTENTIAL).
 FT METAL 36 36 HG(2+) (POTENTIAL).
 SQ SEQUENCE 91 AA: 9611 MW: E1298EBCBCE2FBFD4 CRC64:
 Query Match 16.2%; Score 46; DB 1; Length 91;
 Best Local Similarity 40.9%; Pred. No. 32;
 Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 QY 11 PLWAMFORARVSSDYCATPP 32

```

Db      16 PWWAATQTVTLVSGMTASCSP 37

RESULT 4
ACQO_CHICK
ID ACQO_CHICK STANDARD: PRT: 98 AA.
AC P07032;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Acylphosphatase, organ-common type isozyme (EC 3.6.1.7)
DE (Acylphosphate phosphohydrolase) (Isozyme CH2).
OS Gallus gallus (Chicken).
OC Archaeoptera; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Gallus.
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC NCBI_TaxID=9031;
RN [1]
PE SEQUENCE.
RA TISSUE=Muscle;
RA MEDLINE=88139274; PubMed=2830254;
RT Ohba Y., Minowa O., Mizuno Y., Shiohawa H.;
RT "The primary structure of chicken muscle acylphosphatase isozyme
RT Ch2."
RL J. Biochem. 102:1221-1229(1987).
CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.
CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O -> a fatty acid anion
CC + phosphate.
CC -1- TISSUE SPECIFICITY: ORGAN-COMMON TYPE ISOZYME IS FOUND IN MANY
CC DIFFERENT TISSUES.
CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
DR PIR: A1513; OPC42.
DR HSP: P41500; ZACT.
DR InterPro: IPR001792; Acylphosphatase.
DR Pfam: PF00708; Acylphosphatase; 1.
DR PRINTS: PR00112; ACYLPHPTASE.
DR ProDom: PD001884; Acylphosphatase; 1.
DR PROSITE: PS00150; ACYLPHOSPHATASE_2; 1.
DR PROSITE: PS00151; ACYLPHOSPHATASE_1; 1.
KM Hydrolase; Acetylation; Multigene family.
FT MOD_RES 1 1 ACETYLATION.
FT ACT_SITE 23 23 ACETYLATION.
FT ACT_SITE 41 41 POTENTIAL.
SO SEQUENCE 98 AA; 11019 MW; 1AC35C26218A7E5 CRC64;

Query Match 15.8%; Score 45; DB 1; Length 98;
Best Local Similarity 26.1%; Pred. No. 46;
Matches 12; Conservative 9; Mismatches 13; Indels 12; Gaps 2;

8 RARPLAMWFOR-----ARVSSSDVTCATPPROGDRALREADFO 48
Db 57 RVRLEQWLRKIGSPQSRISREFT-----NEKEIALHETTDQ 95

RESULT 5
YKRL_CAEEL
ID YKRL_CAEEL STANDARD: PRT: 85 AA.
AC P34279;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein CO2F5.2 in chromosome III.
DE CO2F5.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea;
OC Rhabdilitidae; Peleodierinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
PE SEQUENCE FROM N.A.
RP STRAIN=Bristol N2.
RX MEDLINE=94150718; PubMed=7906398;
RX Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,

```

```

RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kersey J., Kirsten J., Laisner N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen M.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L14745; AAA27917.1; -.
DR PIR: S44604; S44604.
DR WormPep: C02F5.2; CE00038.
KW Hypothetical protein.
SO SEQUENCE 85 AA; 9425 MW; 10D565D0551A2CCC CRC64;

Query Match 15.5%; Score 44; DB 1; Length 85;
Best Local Similarity 32.1%; Pred. No. 53;
Matches 9; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 22 SSSDVTCAATPPROGDRALREADFO 49
Db 11 SKDPSISLPPSSGEGALQNTSEEMNA 38

RESULT 6
POL_HV1J3
ID POL_HV1J3 STANDARD: PRT: 100 AA.
AC P12498;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POL polypeptide [contains: Protease (Retropesin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]
DE (Fragment).
DE POL.
GN POL.
OS Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11694;
RN [1]
PE SEQUENCE FROM N.A.
RP MEDLINE=89352108; PubMed=2669897;
RX Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;
RA "Nucleotide sequences of gag and env genes of a Japanese isolate of
RT HIV-1 and their expression in bacteria."
RL AIDS Res. Hum. Retroviruses 5:411-419(1989).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + (DNA)(N).
CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE
CC DETERMINED.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M21137; AAB03523.1; -.
DR HIV; M21137; POLSUI3.
DR MEROPS; A02_001; -.
DR InterPro: IPR001995; Asprotease_rtfv.
DR InterPro: IPR001969; Asprotease_site.
DR Pfam; PF00077; rvp_1.
DR PROSITE; PS00141; ASP_PROTASE; 1.
DR PROSITE; PS01179; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease.
KM Nucleic Acid Transferase; RNA-directed DNA polymerase.
FT CHAIN 57 >100
FT ACT_SITE 81 81
FT NON_TER 100 100 BY SIMILARITY.
SO SEQUENCE 100 AA; 11208 MW; C90A958DEF327254 CRC64;
-----
Very Match 15.0%; Score 42.5; DB 1; Length 100;
Local Similarity 42.4%; Pred. No. 98;
Matches 14; Conservative 2; Mismatches 12; Indels 5; Gaps 1;
-----
OY 18 RARVSSDYTCATPPER-----QGRDLALRLRA 45
DB 12 KAREFSSBOTRANSPSGELAYGHDNNPLSEA 44
:| | :| | :| | :| | :| | :| |
RESULT 7
MERP_PSEAE STANDARD: PRT; 91 AA.
ID MERP_PSEAE
AC P04131.
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mercurotic transport protein periplasmic component precursor
DE (Periplasmic mercury ion binding protein) (Mercury scavenger protein).
GN MERP.
OS Pseudomonas aeruginosa.
OG Plasmid pVS1.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxId=287;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=Tn501; PubMed=6091128;
RX MEDLINE=85014891; Pubmed=6091128;
RY Mistra T.K., Brown N.L., Filtzinger D.C., Pridmore R.D., Barnes W.M.,
RZ Haberstroh L., Silver S.;
S "Mercurotic ion-resistance operons of plasmid R100 and transposon
Tn501: the beginning of the operon including the regulatory region
and the first two structural genes";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5975-5979(1984).
SC -I- FUNCTION: MERCURY SCAVENGER THAT SPECIFICALLY BINDS TO ONE HG(2+)
ION AND WHICH PASSES IT TO THE MERCURIC REDUCTASE (MERA) VIA THE
MERT PROTEIN.
CC -I- SUBUNIT: MONOMER.
CC -I- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC -I- SIMILARITY: CONTAINS 1 HMA DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outpost at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commerce
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
-----
EMBL; Z00027; CAA77323.1; -
EMBL; R02503; AAA2743.1; -
PIR; A03557; RGPSHA.
HSSP; P04129; IAFU.
InterPro; IPR001802; HG_scaevenger
```

```

DR InterPro:IPR001934; HeavyMe_transp.
DR Pfam: PF00403; HMA_1.
DR PRINTS: PR00946; HGSCAVENGER.
DR PROSITE: PS01047; HMA_1; 1.
DR PROSITE: PS00846; HMA_2; 1.
DR Transports: Mercuric resistance; Periplasmic; Metal-binding; Signal;
KW Transposable element; plasmid.
FT SIGNAL 1 19
FT CHAIN 20 91
FT DOMAIN 23 89
FT METAL 33 33
FT METAL 36 36
FT METAL 36 36
SO SEQUENCE 91 AA; 9491 MW; 6D6DB8B5FCA20CE CRC64;

Query Match 14.8%; Score 42; DB 1; Length 91;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 11 PLWAFQARVSSSVTCATPP 32
Db 16 PYWATQFTVLTSLVPMGTCACP 37

RESULT 8
MERP_PSEFL STANDARD: PRT: 91 AA.
ID MERP_PSEFL
AC Q51770:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mercuric transport protein periplasmic component precursor
DE (Periplasmic mercury ion binding protein) (Mercury scavenger protein).
GN MERP.
OS Pseudomonas fluorescens.
OG Pseudomonadaceae;
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE:94341572; PubMed-8063107;
RA Hobman J., Kholodil G., Nikiforov V., Ritchie D.A., Strike P.,
RA Yurileva O.;
RT "The sequence of the mer operon of pMER327/419 and transposon ends of
RT pMER327/419, 330 and 05.";
RL Gene 146:73-78(1994).
CC -1- FUNCTION: MERCURY SCAVENGER THAT SPECIFICALLY BINDS TO ONE HG(2+)
CC ION AND WHICH PASSES IT TO THE MERCURIC REDUCTASE (MERA) VIA THE
CC MERT PROTEIN.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X73112; CNA51540.1; -.
DR HSSP: P04129; IAPF.
DR InterPro: IPR001802; HG_scaevenger.
DR InterPro: IPR001934; HeavyMe_transp.
DR Pfam: PF00403; HMA_1.
DR PRINTS: PR00946; HGSCAVENGER.
DR PROSITE: PS01047; HMA_1; 1.
DR PROSITE: PS00846; HMA_2; 1.
KW Transport; Mercuric resistance; Periplasmic; Metal-binding; Signal;
KW Plasmid.
FT SIGNAL 1 19
FT CHAIN 20 91
FT DOMAIN 23 89
FT METAL 33 33
FT METAL 36 36
FT METAL 36 36
SO SEQUENCE 91 AA; 9491 MW; 6D6DB8B5FCA20CE CRC64;

```

FT CHAIN 20 91 MERCURIC TRANSPORT PROTEIN PERIPLASMIC
 FT DOMAIN 23 89 COMPONENT.
 FT METAL 33 33 HMA.
 FT METAL 36 36 HG(2+) (POTENTIAL).
 SO SEQUENCE 91 AA: 9487 MW: 2D1C3646BA83BE46 CRC64;

Query Match 14.8% Score 42; DB 1; Length 91;
 Best Local Similarity 36.4%; Pred. No. 1e+02;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 11 PLWAMFORARVSSSDVTCATPP 32
 DB 16 PVMATQTWTLVPGMTCSACP 37

RESULT 9
 MERP_SERMA STANDARD: PRT; 91 AA.
 AC P13113;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Mercuric transport protein periplasmic component precursor
 GN (Periplasmic mercury ion binding protein) (Mercury scavenger protein).
 OS Serratia marcescens.
 OC Plasmid pD01358.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89327136; PubMed=2666393;
 RA Nuclefior G., Chu L., Silver S., Miera T.K.;
 RA "Mercury operon regulation by the merp gene of the organomercurial
 RT resistance system of plasmid pD01358.";
 RT J. Bacteriol. 171:4241-4247(1989).
 CC -1- FUNCTION: MERCURY SCAVENGER THAT SPECIFICALLY BINDS TO ONE HG(2+)
 CC ION AND WHICH PASSES IT TO THE MERCURIC REDUCTASE (MERA) VIA THE
 CC MERT PROTEIN.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic (Probable).
 CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M24940; AAA98223.1; -.
 DR PIR: C33858; C33858.
 DR HSP: P04129; IAFV.
 DR InterPro: IPR001802; HG.scavenger.
 DR InterPro: IPR001934; Heavy_mer_transpt.
 DR Pfam: PF00403; HMA_1.
 DR PRINTS: PR00946; HGSCAVENGER.
 DR PROSITE: PS01047; HMA_1; 1.
 DR PROSITE: PS50846; HMA_2; 1.
 DR Transport: Mercuric resistance; Periplasmic; Metal-binding; Signal;
 KW Plasmid.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 91 MERCURIC TRANSPORT PROTEIN PERIPLASMIC
 FT DOMAIN 23 89 COMPONENT.
 FT METAL 33 33 HMA.
 FT METAL 36 36 HG(2+) (POTENTIAL).
 SO SEQUENCE 91 AA: 9548 MW: 21E80D79E9795069 CRC64;

Query Match 14.8% Score 42; DB 1; Length 91;

Best Local Similarity 36.4%; Pred. No. 1e+02;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 11 PLWAMFORARVSSSDVTCATPP 32
 DB 16 PVMATQTWTLVPGMTCSACP 37

RESULT 10
 MERP_SHIFL STANDARD: PRT; 91 AA.
 AC P04129; P07042;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 01-NOV-1986 (Rel. 03, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mercuric transport protein periplasmic component precursor
 DE (Periplasmic mercury ion binding protein) (Mercury scavenger protein).
 GN Shigella flexneri.
 OS Shigella flexneri.
 OC Plasmid IncFII NR1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=85014891; PubMed=6091128;
 RA Miera T.K., Brown N.L., Fritzlinger D.C., Pridmore R.D., Barnes W.M.,
 RA Habersiron L., Silver S.;
 RA "Mercuric ion-resistance operons of plasmid R100 and transposon
 RT Tn501: the beginning of the operon including the regulatory region
 RT and the first two structural genes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 81:5975-5979(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA TRANSDOSON=TN21;
 RA MEDLINE=85159407; PubMed=6530603;
 RA Barthelemy P., Gilbert P., Jackson W.J., Jones C.S., Summers A.O.,
 RA Wisdom S.;
 RA "The DNA sequence of the mercury resistance operon of the IncFII
 RT plasmid NR1.";
 RT J. Mol. Appl. Genet. 2:601-619(1984).
 RN [3]
 RP STRUCTURE BY NMR.
 RA MEDLINE=97332449; PubMed=9188683;
 RA Steele R.A., Opella S.J.;
 RA "Structures of the reduced and mercury-bound forms of Merp, the
 RT periplasmic protein from the bacterial mercury detoxification
 RT system.";
 RL Biochemistry 36:6885-6895(1997).
 RN [4]
 RP STRUCTURE BY NMR.
 RA MEDLINE=98313265; PubMed=9649312;
 RA Qian H., Sahlihan L., Eriksson P.O., Hambræus C., Edlund U.,
 RA Selstam I.;
 RA "NMR solution structure of the oxidized form of Merp, a mercuric ion
 RT binding protein involved in bacterial mercuric ion resistance.";
 RL Biochemistry 37:9316-9322(1998).
 CC -1- FUNCTION: MERCURY SCAVENGER THAT SPECIFICALLY BINDS TO ONE HG(2+)
 CC ION AND WHICH PASSES IT TO THE MERCURIC REDUCTASE (MERA) VIA THE
 CC MERT PROTEIN.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic (Probable).
 CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J01730; AAA92262.1; -.


```

OS Escherichia coli.
OG Plasmid Incp-alpha RK2.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
OX Escherichia.
RN NCBI_TaxID=562.
RP [1]
RC SEQUENCE FROM N.A.
RA STRAIN=K12.
RX MEDLINE=88262548; PubMed=2838814;
RA Thomas C.M., Ibbotson J.P., Wang N., Smith C.A., Tipling R.,
RT "Gene regulation on broad host range plasmid RK2: Identification of
RT three novel operators whose transcription is repressed by both Kora and
RL Nucleic Acids Res. 16:5345-5359(1988).
RP [2]
RC SEQUENCE FROM N.A.
RA MEDLINE=93352413; PubMed=8349548;
RA Kornacki J.A., Chang C.-H., Figueiredo D.H.,
RT "Kil-kor regulation of promiscuous plasmid RK2: structure, products, and
RT regulation of two operators that constitute the kilE locus."
RN J. Bacteriol. 175:5078-5090(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07248; CAA30234.1; -
DR EMBL; L16919; AAA92766.1; -
DR PIR; S00971; S00971.
KW Plasmid; DNA-binding; Transcription regulation.
FT DNA BIND 9 H+H MOTIF (BY SIMILARITY).
SO SEQUENCE 71 AA; 7605 MW; BDD9D998ABDAE04C CRC64;

Query Match 14.4%; Score 41; DB 1; Length 71;
Best Local Similarity 52.4%; Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

OY 28 CATPPER---OGRDLRALRE 44
DB 44 CITPEERORECTLLAALAKO 64

RESULT 14
RN BPMD2
OG VG30_BPMD2 STANDARD: PRT: 88 AA.
OG 0642224:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gene 30 protein (GP30).
GN 30.
OS Mycobacteriophage D29.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OX unclassified Siphoviridae.
RN NCBI_TaxID=28369;
RN [1]
RC SEQUENCE FROM N.A.
RA MEDLINE=96300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: Implications for phage
RT evolution."
RN J. Mol. Biol. 279:143-164(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022214; AAC18471.1; -
SO SEQUENCE 88 AA; 10391 MW; 52071A2E676E5BCE CRC64;

```

```

Query Match 14.1%; Score 40; DB 1; Length 88;
Best Local Similarity 26.2%; Pred. No. 1.8e+02;
Matches 17; Conservative 7; Mismatches 19; Indels 22; Gaps 4;

```

```

OY 1 NPW-----ACDCRRP--LMAFORARVSS-----SDVTCATPPERGRDLR 40
DB 9 NPWETALFFAVFCVLPALPFFMKIKIDSOVNSHDENLRDEITRGFEVNE--DIR 66
OY 41 ALREA 45
DB 67 LHEA 71

```

RESULT 15

```

ID VNS1_IACAO STANDARD: PRT: 90 AA.
AC P26148;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonstructural protein NS1 (Fragment).
OS Influenza A virus (strain A/Camel/Mongolia/82).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94068928; PubMed=8249279;
RA Yamniva S.S., Mandler J., Behk-Ochir Z.H., Dachtzeren P., Ludwig S.,
RA Lvov D.K., Scholtissek C.;
RT "A reassortant H1N1 influenza A virus caused fatal epizootics among
RT camels in Mongolia."
RL Virology 197:558-563(1993).
CC -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF SEGMENT 8.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M73977; AAI16908.2; -
DR HSSP; P03495; INS1.
DR InterPro; IPR000256; Flu_NS1.
DR Pfam; PF00600; Flu_NS1.1.
DR ProDom; PD000613; Flu_NS1.1.
KW Nonstructural protein; Alternative splicing.
FT NON_TER 1
FT NON_TER 1
SO SEQUENCE 90 AA; 10396 MW; 5F11F851E5CE0037 CRC64;

```

```

Query Match 14.1%; Score 40; DB 1; Length 90;
Best Local Similarity 34.5%; Pred. No. 1.8e+02;
Matches 10; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

```

```

OY 15 WFORARVSSDVTCATPPERGRDLRALR 43
DB 4 WVKRRVADQELGDAPFLDRLRDKSLR 32

```

Search completed: April 4, 2003, 08:25:23
Job time : 12.3704 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 08:21:41 ; Search time 18.889 Seconds
(without alignments)
254.473 Million cell updates/sec

Title: US-09-972-546-2_COPY_261_310

Perfect score: 284
Sequence: 1 NPMACDCRRARPLMWFQRR.....PPERGRDLRALREADFQAC 50

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 283224 seqs, 96134422 residues

1 number of hits satisfying chosen parameters: 28758

Minimum DB seq length: 40
Maximum DB seq length: 100

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	57	20.1	100	2	A46308 E3 class 1 protein
2	53	18.7	100	2	S30171 mercuric ion trans
3	50.5	17.8	87	2	T30718 hypothetical prote
4	49.5	17.4	88	2	AD3180 agrobacterium viru
5	48.5	17.1	90	2	JE0210 proteinase (EC 3.4
6	45	15.8	62	2	B82685 hypothetical prote
7	45	15.8	98	1	OPCH2 acylphosphatase (E
8	44	15.5	85	2	S44604 CO2F5.2 protein -
9	43	15.1	86	2	A00924 acetylhydroxy acid
10	42	14.8	65	2	S78743 protein YCR099c -
11	42	14.8	79	2	A87553 hypothetical prote
12	42	14.8	85	2	AB0475 acetolactate synth
13	42	14.8	91	1	REEBHD mercuric resistanc
14	42	14.8	91	1	SQ9524 mercuric resistanc
15	42	14.8	91	1	RQPSHA mercuric resistanc
16	42	14.8	91	1	C3858A mercuric resistanc
17	42	14.8	91	1	S70144 mercuric resistanc
18	42	14.8	97	2	A60127 mercuric resistanc
19	41.5	14.6	80	2	T65235 homeotic protein a
20	41	14.4	58	2	AF2538 testicular luteinl
21	41	14.4	71	2	C53306 hypothetical prote
22	41	14.4	72	2	A93351 Kieb protein - pla
23	40.3	14.3	57	2	C82501 protein (imported
24	40.5	14.3	93	2	AB0740 hypothetical prote
25	40.5	14.3	97	2	C87381 probable excisiona
26	40	14.1	66	2	AF2423 hypothetical prote
27	40	14.1	86	2	E69071 hypothetical prote
28	40	14.1	88	2	D72803 gp30 protein - Myc
29	40	14.1	94	2	F84193 hypothetical prote

ALIGNMENTS

```

30      40      14.1      98      2      S33482      hypothetical prote
31      39.5      13.9      72      2      I50024      sky-related sequen
32      39.5      13.9      73      2      AC3365      hypothetical prote
33      39.5      13.9      87      2      T47209      ccg-4 protein 2 (l
34      39.5      13.9      90      2      S19597      transcription fact
35      39.5      13.9      94      2      S28802      hypothetical prote
36      39.5      13.9      95      2      C66685      prophage p12 prote
37      39.5      13.9      95      2      B86757      holin protein - less
38      39.5      13.9      95      2      T13259      sox protein - less
39      39      13.7      56      2      S33895      hypothetical prote
40      39      13.7      71      2      H95299      homeotic protein 2
41      39      13.7      90      2      JC1163      hypothetical prote
42      39      13.7      97      2      AD3030      E3 protein - bovin
43      39      13.7      98      1      M3MLB4      hypothetical prote
44      39      13.7      99      2      T10554      excisionase - phag
45      38.5      13.6      65      1      RSBPX8

RESULT 1
A46308
E3 class 1 protein - mouse adenovirus 1
C:Species: Mastadenovirus mus1 (mouse adenovirus 1)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C:Accession: A46308
R:Beard, C.W.; Ball, A.O.; Wooley, E.H.; Spindler, K.R.
Virology 175, 81-90, 1990
A:Title: Transcription mapping of mouse adenovirus type 1 early region 3.
A:Reference number: A46308; MUID:90177241; PMID:2137954
A:Accession: A46308
A:Molecule type: DNA
A:Residues: 1-100 <BEA>
A:Cross-references: GB:M33995; NID:g209915; PIDN:AAA42494.1; PID:g209917
C:Genetics:
A:Introns: 15/1; 72/2
C:Superfamily: mouse adenovirus early E3 protein

Query Match      20.1%      Score 57; DB 2; Length 100;
Best Local Similarity 29.2%      Pred. No. 5.2;
Matches 14; Conservative 6; Mismatches 22; Indels 6; Gaps 1;

QY      2      PMACDCRRARPLM-----AMFORAVSSSDVTCATPPERGRDLRALR 43
Db      43      PSQCPCPASPPMTNSVTSTFAQTKMENSROYCPVPSBSTRGKNAVR 90

RESULT 2
S30171
mercuric ion transport protein merT - Streptomyces lividans
C:Species: Streptomyces lividans
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Oct-1999
C:Accession: S30171; S23611
R:Sedlmeyer, R.; Altenbuchner, J.
Mol. Gen. Genet. 236, 76-85, 1992
A:Title: Cloning and DNA sequence analysis of the mercury resistance genes of Strepto
A:Reference number: S30168; MUID:93156667; PMID:1494553
A:Accession: S30171
A:Molecule type: DNA
A:Residues: 1-100 <ALT>
A:Cross-references: EMBL:X65467; NID:g47180; PIDN:CAA46463.1; PID:g47184
A:Experimental source: strain 1326
C:Genetics:
A:Gene: merT
A:Function:
A:Description: Involved in transport of mercury ions from periplasm into cytoplasm; p
A:Pathway: mercury resistance
A:Note: mercury resistance operon
C:Keywords: binding protein-dependent transport system; mercury transport; transmembr
F:15-31/Domain: transmembrane #status predicted <TM1>
F:36-52/Domain: transmembrane #status predicted <TM2>
F:55-71/Domain: transmembrane #status predicted <TM3>

```


C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Aug-1994
C:Accession: A41513

R:Ohba, Y.; Minowa, O.; Mizuno, Y.; Shiohawa, H.
J. Biochem. 102, 1221-1229, 1987

A:Title: The primary structure of chicken muscle acetylphosphatase isozyme Ch2.
A:Reference number: A41513; MUID:88139274; PMID:2650254

A:Accession: A41513

A:Molecule type: protein
A:Residues: 1-98 <OH>
C:Superfamily: acylphosphatase
C:Keywords: acetylated amino end; hydrolase
F:/Modified site: acetylated amino end (Ala) #status experimental

Query Match 15.8%; Score 45; DB 1; Length 98;
Best Local Similarity 26.1%; Pred. No. 1.6e+02;
Matches 12; Conservative 9; Mismatches 13; Indels 12; Gaps 2;

OY 8 RAPPLMAMFOR----ARVSSDYCATPPEROGRDLALREADFOQ 48
| | | | | : | | : | | |
57 RRELQELMRKIGSPQSIRISRAEFT-----NEKEIALLEHFDFO 95

RESULT 8
S44604

C02P5.2 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-1997
C:Accession: S44604
R:Anderson, K.
submitted to the EMBL Data Library, May 1993
A:Description: Sequence of the C. elegans cosmid C02P5.
A:Reference number: S44603
A:Accession: S44604
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-85 <AND>
A:Cross-references: EMBL:L14745; NID:g289607; PID:g289609
C:Genetics:
A:introns: 19/1; 54/3

Query Match 15.5%; Score 44; DB 2; Length 85;
Best Local Similarity 32.1%; Pred. No. 1.8e+02;
Matches 9; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 22 SSDDVCATPPEROGRDLALREADFOQ 49
| | | | | : | | : | | |
PH 11 SKSDTSLPSSTGALONTTESEWNA 38

RESULT 9
AE0924

acetylhydroxy acid synthase II, small chain [Imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE0924
R:Parthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, M.T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; ...
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov...
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0924
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD09415.1; PID:g16504533; GSPDB:GN00176
C:Genetics:
A:Gene: STY3655

Query Match 15.1%; Score 43; DB 2; Length 86;

```

Best Local Similarity 38.5%; Pred No. 2.5e+02;
Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0
QY      25 DYTCA TPPE RGRG DRLALREAD FQAC 50
      ::  ||  ||  ||  ||  ||  ||  ||
Db       6 NVSARFPETLERVLRVRRHGFQVC 31

RESULT 10
578743
protein YCR099c - yeast (Saccharomyces cerevisiae)
C.Species: Saccharomyces cerevisiae
C.Date: 15-Jan-1999 #sequence_revision 15-Jan-1999 #text_change 15-Jan-1999
C.Accession: S78743
R.Grievell, L.A.; de Haan, M.; Maat, M.J.
submitted to the Protein Sequence Database, March 1992
A.Reference number: S19412
A.Accession: S78743
A.Status: Preliminary
A.Molecule type: DNA
A.Residues: 1-65 <GR1>
A.Cross-references: EMBL:X59720; MIPS:YCR102w-a
C.Genetics:
A:Map position: 3R

```

Query Match	14.8%	Score 42;	DB 2;	Length 65;
Best Local Similarity	50.0%	Pred. No. 2.5e+02;		
Matches	7;	Conservative	4;	Mismatches 3;
				Indels 0;
				Gaps 0;
QY	12	LMWFGARARSSD	25	
	11	111111111111		
Db	48	LMKTIQEQQLSSND	61	

RESULT 11
A87553
hypothetical protein CC2450 [Imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: A87553
R:Miernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heldberg,
n, J.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Hart, D.H.; Ko-
b, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87553
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-79 <SMO>
A:Cross-references: GB:AE005673; MID:q13423997; PIDN:AAK24421.J; GSPDB:GN00148
A:Gene: CC2450

```

Query Match          14.8%  Score 42;  DB 2;  Length 79;
Best Local Similarity 50.0%  Pred. No. 3e+02;
Matches 14;  Conservative 1;  Mismatches 9;  Indels 4;  Caps 2;

QY      27  TCATP--PEROGRDRLALREAD--FQAC 50
          ||| | | | | | | | | |
Db       10  TCADPRDGRDRNRHRCGLADRADQAC 37

RESULT 12
AB0475
acetoillactate synthase (EC 4.1.3.18) isozyme II small chain [imported] [imported] - Ye
C:Species: Versinia pectis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB0475
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.
delen-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
ill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
ature 413, 523-527, 2001

```

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AB0475
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-85 <KUD>
 A:Cross-references: GB:AL590842; PIDN:CAC93366.1; PID:g15981812; GSPDB:GN00175
 C:Genetics:
 A:Gene: *lyvM*
 C:Keywords: carbon-carbon lyase; oxo-acid-lyase

Query Match 14.8%; Score 42; DB 2; Length 85;
 Best Local Similarity 47.4%; Pred. No. 3.4e+02;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 32 PERGGRDLRALREADFOAC 50
 |||||
 Db 14 PEMERVLVYVHRHGFQVC 32

LT 13

mercuric resistance operon regulatory protein merp precursor - *Shigella flexneri* plasmid
 N:Alternate names: mercuric-ion-binding protein; mercury scavenger protein
 C:Species: *Shigella flexneri*
 C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999
 C:Accession: A03556
 R:Mirza, T.K.; Brown, N.L.; Fritzing, D.C.; Priddy, R.D.; Barnes, W.M.; Haberstroch, Proc. Natl. Acad. Sci. U.S.A. 81, 5975-5979, 1984
 A:Title: Mercuric ion-resistance operons of plasmid R100 and transposon Tn501: the begin
 A:Reference number: A03556; MUID:85014891; PMID:6091128
 A:Accession: A03556
 A:Molecule type: DNA
 A:Residues: 1-91 <MIS>
 A:Cross-references: GB:J01730; NID:g151742; PIDN:AAA92262.1; PID:g151751
 C:Genetics:
 A:Gene: merp; merc
 A:Genome: plasmid
 C:Superfamily: mercuric resistance operon regulatory protein; heavy-metal-associated hom
 C:Keywords: mercury transport; metal binding; periplasmic space
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-91/Product: mercuric resistance operon regulatory protein merp #status predicted <M
 F:28-57/Domain: heavy-metal-associated homology <HMA>
 F:33,36/Binding site: mercury (Cys) #status predicted

Query Match 14.8%; Score 42; DB 1; Length 91;
 Best Local Similarity 36.4%; Pred. No. 3.4e+02;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 11 PLWAMFORARVSSDYTCATPP 32
 |||||
 Db 16 PVWATQTVTILAVPGMTCAACP 37

RESULT 14

S09524
 mercuric resistance operon regulatory protein precursor - plasmid NR1
 N:Alternate names: mercuric-ion-binding protein; mercury scavenger protein
 C:Species: plasmid NR1
 C:Date: 16-Oct-1998 #sequence_revision 16-Oct-1998 #text_change 16-Jul-1999
 C:Accession: S09524
 R:Barthelemy, P.; Gilbert, P.; Jackson, W.J.; Jones, C.S.; Summers, A.O.; Wisdom, S. J. Mol. Appl. Genet. 2, 601-619, 1984
 A:Title: The DNA sequence of the mercury resistance operon of the IncFII plasmid NR1.
 A:Reference number: S07447; MUID:85154407; PMID:6530603
 A:Accession: S09524
 A:Molecule type: DNA
 A:Residues: 1-91 <BAR>
 A:Cross-references: EMBL:K03089; NID:g150389; PIDN:AAB59076.1; PID:g150394
 C:Genetics:
 A:Gene: merTB
 A:Genome: plasmid
 C:Superfamily: mercuric resistance operon regulatory protein; heavy-metal-associated hom

C:Keywords: mercury transport; metal binding; periplasmic space
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-91/Product: mercuric resistance operon regulatory protein #status predicted <MAT
 F:28-57/Domain: heavy-metal-associated homology <HMA>
 F:33,36/Binding site: mercury (Cys) #status predicted

Query Match 14.8%; Score 42; DB 1; Length 91;
 Best Local Similarity 36.4%; Pred. No. 3.4e+02;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 11 PLWAMFORARVSSDYTCATPP 32
 |||||
 Db 16 PVWATQTVTILAVPGMTCAACP 37

RESULT 15

RGPSHA
 mercuric resistance operon regulatory protein merp precursor - *Pseudomonas aeruginosa*
 N:Alternate names: mercuric-ion-binding protein; mercury scavenger protein
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999
 C:Accession: A03557
 R:Mirza, T.K.; Brown, N.L.; Fritzing, D.C.; Priddy, R.D.; Barnes, W.M.; Haberstroch, Proc. Natl. Acad. Sci. U.S.A. 81, 5975-5979, 1984
 A:Title: Mercuric ion-resistance operons of plasmid R100 and transposon Tn501: the be
 A:Reference number: A03556; MUID:85014891; PMID:6091128
 A:Accession: A03557
 A:Molecule type: DNA
 A:Residues: 1-91 <MIS>
 A:Cross-references: GB:K02503; NID:g154897; PIDN:AAA27434.1; PID:g154900
 C:Genetics:
 A:Gene: merp; merc
 A:Genome: plasmid
 C:Superfamily: mercuric resistance operon regulatory protein; heavy-metal-associated
 C:Keywords: mercury transport; metal binding; periplasmic space
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-91/Product: mercuric resistance operon regulatory protein merp #status predicted
 F:28-57/Domain: heavy-metal-associated homology <HMA>
 F:33,36/Binding site: mercury (Cys) #status predicted

Query Match 14.8%; Score 42; DB 1; Length 91;
 Best Local Similarity 36.4%; Pred. No. 3.4e+02;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 11 PLWAMFORARVSSDYTCATPP 32
 |||||
 Db 16 PVWATQTVTILAVPGMTCAACP 37

Search completed: April 4, 2003, 08:24:48
 Job time : 21.8889 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 12:08:21 ; Search time 19 seconds

(without alignments)
1351.428 Million cell updates/sec

Title: US-09-972-546-2

Sequence: 1 MBLGRLRLQAPASACLIM.....LSAGLPPLCLLLVPHNL 420

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

1 number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications-AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	808	36.3	473	US-09-978-295A-400	Sequence 400, App
2	808	36.3	473	US-09-978-697-400	Sequence 400, App
3	808	36.3	473	US-09-978-192A-400	Sequence 400, App
4	808	36.3	473	US-09-999-832A-400	Sequence 400, App
5	808	36.3	473	US-09-978-189-400	Sequence 382, App
6	808	36.3	473	US-10-028-072-382	Sequence 382, App
7	808	36.3	473	US-10-121-049-382	Sequence 382, App
8	808	36.3	473	US-10-123-904-382	Sequence 382, App
9	808	36.3	473	US-10-140-470-382	Sequence 382, App
10	808	36.3	473	US-10-175-746-382	Sequence 382, App
11	808	36.3	473	US-10-176-921-382	Sequence 382, App
12	808	36.3	473	US-10-227-884-128	Sequence 128, App
13	808	36.3	473	US-10-137-865-382	Sequence 382, App
14	808	36.3	473	US-10-140-474-382	Sequence 382, App
15	808	36.3	473	US-10-142-431-382	Sequence 382, App
16	808	36.3	473	US-10-143-114-382	Sequence 382, App
17	808	36.3	473	US-10-230-163-128	Sequence 128, App
18	808	36.3	473	US-10-140-002-382	Sequence 382, App
19	808	36.3	473	US-10-140-002-382	Sequence 382, App

20	808	36.3	473	US-09-978-608A-400	Sequence 400, App
21	808	36.3	473	US-10-142-419-382	Sequence 382, App
22	808	36.3	473	US-10-218-631-128	Sequence 128, App
23	808	36.3	473	US-10-230-338-128	Sequence 128, App
24	808	36.3	473	US-09-978-191A-400	Sequence 400, App
25	808	36.3	473	US-09-978-403A-400	Sequence 400, App
26	808	36.3	473	US-09-978-564A-400	Sequence 400, App
27	808	36.3	473	US-09-978-564A-400	Sequence 400, App
28	808	36.3	473	US-10-017-081A-400	Sequence 400, App
29	808	36.3	473	US-10-123-262-382	Sequence 382, App
30	808	36.3	473	US-10-142-423-382	Sequence 382, App
31	808	36.3	473	US-10-230-414-128	Sequence 128, App
32	808	36.3	473	US-09-978-824-400	Sequence 400, App
33	808	36.3	473	US-09-981-915A-400	Sequence 400, App
34	808	36.3	473	US-09-999-832A-400	Sequence 400, App
35	808	36.3	473	US-10-121-050-382	Sequence 382, App
36	808	36.3	473	US-10-141-755-382	Sequence 382, App
37	808	36.3	473	US-10-167-749-400	Sequence 400, App
38	808	36.3	473	US-09-918-585A-400	Sequence 400, App
39	808	36.3	473	US-10-143-032-382	Sequence 382, App
40	808	36.3	473	US-09-758-140-2	Sequence 400, App
41	808	36.3	473	US-09-893-348-26	Sequence 2, App1
42	808	36.3	473	US-09-972-599A-2	Sequence 26, App1
43	805	35.2	473	US-09-796-658-34	Sequence 34, App1
44	783.5	35.2	380	US-10-012-542-161	Sequence 161, App1
45	775	34.8	473	US-09-758-140-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-09-978-295A-400
Sequence 400, Application US/09978295A
Patent No. US20020156006A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Godowski, Audrey J.
APPLICANT: Grimaldi, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC11
CURRENT APPLICATION NUMBER: US/09/978, 295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249

1	PRIOR FILING DATE: 1997-11-03
2	PRIOR APPLICATION NUMBER: 60/065311
3	PRIOR FILING DATE: 1997-11-13
4	PRIOR APPLICATION NUMBER: 60/066364
5	PRIOR FILING DATE: 1997-11-21
6	PRIOR APPLICATION NUMBER: 60/077450
7	PRIOR FILING DATE: 1998-03-10
8	PRIOR APPLICATION NUMBER: 60/077632
9	PRIOR FILING DATE: 1998-03-11
10	PRIOR APPLICATION NUMBER: 60/077641
11	PRIOR FILING DATE: 1998-03-11
12	PRIOR APPLICATION NUMBER: 60/077649
13	PRIOR FILING DATE: 1998-03-11
14	PRIOR APPLICATION NUMBER: 60/077791
15	PRIOR FILING DATE: 1998-03-12
16	PRIOR APPLICATION NUMBER: 60/078004
17	PRIOR FILING DATE: 1998-03-13
18	PRIOR APPLICATION NUMBER: 60/078886
19	PRIOR FILING DATE: 1998-03-20
20	PRIOR APPLICATION NUMBER: 60/078936
21	PRIOR FILING DATE: 1998-03-20
22	PRIOR APPLICATION NUMBER: 60/078910
23	PRIOR FILING DATE: 1998-03-20
24	PRIOR APPLICATION NUMBER: 60/078939
25	PRIOR FILING DATE: 1998-03-20
26	PRIOR APPLICATION NUMBER: 60/079294
27	PRIOR FILING DATE: 1998-03-25
28	PRIOR APPLICATION NUMBER: 60/079656
29	PRIOR FILING DATE: 1998-03-26
30	PRIOR APPLICATION NUMBER: 60/079664
31	PRIOR FILING DATE: 1998-03-27
32	PRIOR APPLICATION NUMBER: 60/079689
33	PRIOR FILING DATE: 1998-03-27
34	PRIOR APPLICATION NUMBER: 60/079663
35	PRIOR FILING DATE: 1998-03-27
36	PRIOR APPLICATION NUMBER: 60/079728
37	PRIOR FILING DATE: 1998-03-27
38	PRIOR APPLICATION NUMBER: 60/079786
39	PRIOR FILING DATE: 1998-03-27
40	PRIOR APPLICATION NUMBER: 60/079920
41	PRIOR FILING DATE: 1998-03-30
42	PRIOR APPLICATION NUMBER: 60/079923
43	PRIOR FILING DATE: 1998-03-30
44	PRIOR APPLICATION NUMBER: 60/080105
45	PRIOR FILING DATE: 1998-03-31
46	PRIOR APPLICATION NUMBER: 60/080107
47	PRIOR FILING DATE: 1998-03-31
48	PRIOR APPLICATION NUMBER: 60/080165
49	PRIOR FILING DATE: 1998-03-31
50	PRIOR APPLICATION NUMBER: 60/080194
51	PRIOR FILING DATE: 1998-03-31
52	PRIOR APPLICATION NUMBER: 60/080327
53	PRIOR FILING DATE: 1998-04-01
54	PRIOR APPLICATION NUMBER: 60/080328
55	PRIOR FILING DATE: 1998-04-01
56	PRIOR APPLICATION NUMBER: 60/080333
57	PRIOR FILING DATE: 1998-04-01
58	PRIOR APPLICATION NUMBER: 60/080334
59	PRIOR FILING DATE: 1998-04-01
60	PRIOR APPLICATION NUMBER: 60/081070
61	PRIOR FILING DATE: 1998-04-08
62	PRIOR APPLICATION NUMBER: 60/081049
63	PRIOR FILING DATE: 1998-04-08
64	PRIOR APPLICATION NUMBER: 60/081071
65	PRIOR FILING DATE: 1998-04-08
66	PRIOR APPLICATION NUMBER: 60/081195
67	PRIOR FILING DATE: 1998-04-08
68	PRIOR APPLICATION NUMBER: 60/081203
69	PRIOR FILING DATE: 1998-04-09
70	PRIOR APPLICATION NUMBER: 60/081229
71	PRIOR FILING DATE: 1998-04-09
72	PRIOR APPLICATION NUMBER: 60/081955
73	PRIOR FILING DATE: 1998-04-15

1	PRIOR APPLICATION NUMBER: 60/081811
2	PRIOR FILING DATE: 1998-04-15
3	PRIOR APPLICATION NUMBER: 60/081819
4	PRIOR FILING DATE: 1998-04-15
5	PRIOR APPLICATION NUMBER: 60/081952
6	PRIOR FILING DATE: 1998-04-15
7	PRIOR APPLICATION NUMBER: 60/081838
8	PRIOR FILING DATE: 1998-04-15
9	PRIOR APPLICATION NUMBER: 60/082568
10	PRIOR FILING DATE: 1998-04-21
11	PRIOR APPLICATION NUMBER: 60/082569
12	PRIOR FILING DATE: 1998-04-21
13	PRIOR APPLICATION NUMBER: 60/082704
14	PRIOR FILING DATE: 1998-04-22
15	PRIOR APPLICATION NUMBER: 60/082804
16	PRIOR FILING DATE: 1998-04-22
17	PRIOR APPLICATION NUMBER: 60/082700
18	PRIOR FILING DATE: 1998-04-22
19	PRIOR APPLICATION NUMBER: 60/082797
20	PRIOR FILING DATE: 1998-04-22
21	PRIOR APPLICATION NUMBER: 60/082796
22	PRIOR FILING DATE: 1998-04-23
23	PRIOR APPLICATION NUMBER: 60/083336
24	PRIOR FILING DATE: 1998-04-27
25	PRIOR APPLICATION NUMBER: 60/083322
26	PRIOR FILING DATE: 1998-04-28
27	PRIOR APPLICATION NUMBER: 60/083392
28	PRIOR FILING DATE: 1998-04-29
29	PRIOR APPLICATION NUMBER: 60/083495
30	PRIOR FILING DATE: 1998-04-29
31	PRIOR APPLICATION NUMBER: 60/083496
32	PRIOR FILING DATE: 1998-04-29
33	PRIOR APPLICATION NUMBER: 60/083499
34	PRIOR FILING DATE: 1998-04-29
35	PRIOR APPLICATION NUMBER: 60/083545
36	PRIOR FILING DATE: 1998-04-29
37	PRIOR APPLICATION NUMBER: 60/083554
38	PRIOR FILING DATE: 1998-04-29
39	PRIOR APPLICATION NUMBER: 60/083558
40	PRIOR FILING DATE: 1998-04-29
41	PRIOR APPLICATION NUMBER: 60/083559
42	PRIOR FILING DATE: 1998-04-29
43	PRIOR APPLICATION NUMBER: 60/083500
44	PRIOR FILING DATE: 1998-04-29
45	PRIOR APPLICATION NUMBER: 60/083742
46	PRIOR FILING DATE: 1998-04-30
47	PRIOR APPLICATION NUMBER: 60/084366
48	PRIOR FILING DATE: 1998-05-05
49	PRIOR APPLICATION NUMBER: 60/084414
50	PRIOR FILING DATE: 1998-05-06
51	PRIOR APPLICATION NUMBER: 60/084441
52	PRIOR FILING DATE: 1998-05-06
53	PRIOR APPLICATION NUMBER: 60/084637
54	PRIOR FILING DATE: 1998-05-07
55	PRIOR APPLICATION NUMBER: 60/084639
56	PRIOR FILING DATE: 1998-05-07
57	PRIOR APPLICATION NUMBER: 60/084640
58	PRIOR FILING DATE: 1998-05-07
59	PRIOR APPLICATION NUMBER: 60/084588
60	PRIOR FILING DATE: 1998-05-07
61	PRIOR APPLICATION NUMBER: 60/084600
62	PRIOR FILING DATE: 1998-05-07
63	PRIOR APPLICATION NUMBER: 60/084627
64	PRIOR FILING DATE: 1998-05-07
65	PRIOR APPLICATION NUMBER: 60/084643
66	PRIOR FILING DATE: 1998-05-07
67	PRIOR APPLICATION NUMBER: 60/085339
68	PRIOR FILING DATE: 1998-05-13
69	PRIOR APPLICATION NUMBER: 60/085338
70	PRIOR FILING DATE: 1998-05-13
71	PRIOR APPLICATION NUMBER: 60/085323
72	PRIOR FILING DATE: 1998-05-13
73	PRIOR APPLICATION NUMBER: 60/085562


```

: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085700
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085689
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085579
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085580
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085573
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085704
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085697

```

Query Match 36.3%; Score 808; DB 9; Length 473;
 Best Local Similarity 42.2%; Pred. No. 1.1e-57;

Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

```

18 LMLALPLAPGPMCTCTCYSSPP-TVSCQANNFSSVPLSPSTQRLFLQNNLIRTLR 76
15 VMLQAMQVAP-CPGACVCYNBPKYTTSCPOQLQAPVGPVIPAASQRIFLHGNRISHVP 73
77 PGTFGS--NLTLMLFSNNLSTIYPGTFHRLQLELDLGDNRHLRSLEPDTQGLERQ 134
74 AASFRCRNLTLMLHNSVNLARIDAAFTGLALLEQDLSDNQLKSVDPATFHGLGRH 133
135 SLHLRYCQLSSLEGNIFRGVLVLYLYLOENSLHLQDLFLADLANI,SHLFLHGNRLRL 194
134 TLHLDRGLOELPGLFRGLALQLYLYLDNALQALPDDTFRDLGNLTLHLHGNRISV 193
195 TEHVFRLGLSLRLLHGNLQGVHRAFRGLSRLLTLVLFNNSLASLGEALADLPSE 254
194 PEHAFGLSHSLDLHLHGNRAVHHAFRDGLRLTLVLFANLSALPTALAPLRLAQ 253
255 FLRLANPACDCRRAPLWAFQARVSSDYTCATPPEGRGRDLRLREADFOAC---- 310
254 YRLNDNPMVCDRCRAPLWAMIQKFRGSSSEVPCLPQLARLDLRLANLQCAVAT 313
311 -----PPAA-----PFRGSRAR-----GNS 326
314 GRYHPTWTRATDEBPLGLPKCCQPDAAKASVLEPGRPASAGMLKGRVPPGDSPPGNG 373
327 SS-NHL----YGVAAEAPPADPSTLYR-----DLPAEDSRGRG----- 361
374 SGRPHINDSPFGLPQSAAP--PLTAVREGESEPGFPISGRRRRRRCRKNRTRSHCR 431
362 GDAPTEDDYWGCGYGGEDQGEQMCFGAACQAPP 394
432 GOAGS-----GGGCTGDSGSGALPLSLTCLTP 459

```

RESULT 2

US-09-978-697-400
 Sequence 400; Application US/09978697
 Patent No. US20020169284A1

GENERAL INFORMATION:

```

: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Bolstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Oliang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerlisen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J

```

```

: APPLICANT: Kljavin, Iyar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC27
: CURRENT APPLICATION NUMBER: US/09/978,697
: PRIOR APPLICATION NUMBER: 2001-10-16
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078866
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080327

```

PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06

PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 36.3%; Score 808; DB 9; Length 473;
Best Local Similarity 42.2%; Pred. No. 1,1e-57;
Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

QY 18 LMLALPLAASCPMLCTCYSSP-TVSCANNSSVPLSPFTQTLFQNNLIRFLR 76
DB 15 VMLQMGVAAP-CPGACVCYNEPKVTSCPOGLQAVPGVIGASQRIFLGNHISHP 73
QY 77 PGTGGS--NLLTLMFSSNNLSIYPTGFPHLQALBELDGDNRHRSLEPDTFGLERLQ 134
DB 74 AASFRCRNLTTLMLHSNVLARIDAAFTGLALLDQLDLSDAQLRSDVPATFEGHGRHL 133
QY 135 SLHLVRCQLSLPGNIFRGVLSIQLYLYLOENSLHLLODDLFPADLANLSHLFHNRLRL 194
DB 134 TLHLDRGGLQELGPGCFRGIAALQYLYLODNALQALPDDTFRDLGNTLTLFHNRISSV 193
QY 195 TEHVEFGGLSLRLHLHGNRLQGVNHAARGLSLTLFLFNNSLASLGEALADLPISLE 254
DB 194 PERAFRGLHSLDRLLHONRAVHAHFRDLGRMTLYLFFANNLSALTTELAPLRALQ 253
QY 255 FLRLNPNACDCRRAPRLAMFQARVSSDVYCATPPERGDRFLALREADFOAC---- 310
DB 254 YLRLNPNWCCDRAPRLAMWLQKFRGSSSEVPCSLPQRLAGDRFLRLAANDLQCAVAT 313
QY 311 -----PPAA-----PTRGSRAR-----GNS 326
DB 314 GPHYPIWTGRATDEEPLGPKCCPDAAADKASVLEPGRPASAGNALKGRVPGDSFGNG 373
QY 327 SS-NHL-----YGAVALGAPPADPSTLYR-----DLPAEDSRGROG----- 361
DB 374 GGRPHNDSPFGLPGSAEP-PLTAVRPGSRPPGPTTGPARRRGCSKKNRTRSHCRLL 431
QY 362 GDAFTEDDYWGCGEDQGEQMCPGAACQAPP 394
DB 432 GQAGS-----GGGTGDSSEGSALPILTCSLTP 459

RESULT 3
US-09-978-192A-400
Sequence 400, Application US/09978192A
Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496

[illegible]

RESULT 4
US-09-999-832A-400
: Sequence 400, Application US/0999832A
: Publication No. US20020192706A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Flivaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerlitsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James
: APPLICANT: Paonl, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tunas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630PIC63
: CURRENT APPLICATION NUMBER: US/09/999.832A
: CURRENT FILING DATE: 2001-10-24
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641


```
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match      36.3%  Score 808; DB 9; Length 473;
Best Local Similarity 42.2%; Pred. No. 1,1e-57;
Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

OY 18 LLMALALPLAPSCPMCTCYSSP-TVSCQANNFSSVPLSLPSTORLFLONLIRTLR 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 15 VLMQAMQVAP-CGCAVCYNKPKVTSSCPQGGQAVPGVIPAASQRTFLHGNRISHVP 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 77 PGTGGS--NLTLWLFNSNLTITTYGTRHQALEDLGNHRHRSLEPOTFOGLELQ 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 74 AAFRACNLTLMHSHVLRIDAFAFTGLLELDLSDAQAQSVDPATFHLGRHL 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 135 SLHXRQCLSLPGNIFGVLVSLOYLYLOENSLHLHODDLFADLANLSHLFHLGRHL 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 134 TLHRCGLQELGPGFLFGIALQYLYLODNALQALPDITFEDLGNLHLFHLGRH 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 195 TENVRGGLSDRLHLHGRLQGVHRAAFRGLSLRTITYLENNSLASLPGELADLPSE 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 194 PERAFRGLHSDRLHLHONRVHAFHAFRDLGRIMLTLYLFANNLSALPTEALAEFLRLQ 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 255 FLRLNAPACDCRRAPLWAFQBARVSSDYTCATPPEQGRDLRALREADFOAC--- 310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 YRLNDNPWCDCRRAPLWAFQBARVSSDYTCATPPEQGRDLRALREADFOAC--- 310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 311 -----PPAA-----PTRGSRAR-----GNS 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 314 GPHYHWTGRATDEEPLGPKCCOPDAADKASVLEFGRFASAGNALKGRVPPGDSPPGNG 373
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 327 SS-NHL-----YVALEGARPPADPSTLYR-----DLPEDSRGRG----- 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 374 SGPRIHNDSPFTLPQSAP--PLTAVREGSEBPPEFPLSGRRRRPGCCRKMTSHCL 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 362 GDPATEDYWGYGGEDEQGMCPGACQAPP 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 432 GQAGS-----GGGGTGDSGSGALPLTCTSLTP 459
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-09-978-189-400
; Sequence 400, Application US/09978189
; Publication No. US20030004102A1
GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
```

```
; APPLICANT: Wood, William I.
; TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC7
; CURRENT APPLICATION NUMBER: US/09/978,189
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
```

1 PRIOR FILING DATE: 1998-04-08
2 PRIOR APPLICATION NUMBER: 60/081071
3 PRIOR FILING DATE: 1998-04-08
4 PRIOR APPLICATION NUMBER: 60/081195
5 PRIOR FILING DATE: 1998-04-08
6 PRIOR APPLICATION NUMBER: 60/081203
7 PRIOR FILING DATE: 1998-04-09
8 PRIOR APPLICATION NUMBER: 60/081229
9 PRIOR FILING DATE: 1998-04-09
10 PRIOR APPLICATION NUMBER: 60/081955
11 PRIOR FILING DATE: 1998-04-15
12 PRIOR APPLICATION NUMBER: 60/081817
13 PRIOR FILING DATE: 1998-04-15
14 PRIOR APPLICATION NUMBER: 60/081819
15 PRIOR FILING DATE: 1998-04-15
16 PRIOR APPLICATION NUMBER: 60/081952
17 PRIOR FILING DATE: 1998-04-15
18 PRIOR APPLICATION NUMBER: 60/081838
19 PRIOR FILING DATE: 1998-04-15
20 PRIOR APPLICATION NUMBER: 60/082568
21 PRIOR FILING DATE: 1998-04-21
22 PRIOR APPLICATION NUMBER: 60/082569
23 PRIOR FILING DATE: 1998-04-21
24 PRIOR APPLICATION NUMBER: 60/082704
25 PRIOR FILING DATE: 1998-04-22
26 PRIOR APPLICATION NUMBER: 60/082804
27 PRIOR FILING DATE: 1998-04-22
28 PRIOR APPLICATION NUMBER: 60/082700
29 PRIOR FILING DATE: 1998-04-22
30 PRIOR APPLICATION NUMBER: 60/082797
31 PRIOR FILING DATE: 1998-04-22
32 PRIOR APPLICATION NUMBER: 60/082796
33 PRIOR FILING DATE: 1998-04-23
34 PRIOR APPLICATION NUMBER: 60/083336
35 PRIOR FILING DATE: 1998-04-27
36 PRIOR APPLICATION NUMBER: 60/083322
37 PRIOR FILING DATE: 1998-04-28
38 PRIOR APPLICATION NUMBER: 60/083392
39 PRIOR FILING DATE: 1998-04-29
40 PRIOR APPLICATION NUMBER: 60/083495
41 PRIOR FILING DATE: 1998-04-29
42 PRIOR APPLICATION NUMBER: 60/083496
43 PRIOR FILING DATE: 1998-04-29
44 PRIOR APPLICATION NUMBER: 60/083499
45 PRIOR FILING DATE: 1998-04-29
46 PRIOR APPLICATION NUMBER: 60/083545
47 PRIOR FILING DATE: 1998-04-29
48 PRIOR APPLICATION NUMBER: 60/083554
49 PRIOR FILING DATE: 1998-04-29
50 PRIOR APPLICATION NUMBER: 60/083558
51 PRIOR FILING DATE: 1998-04-29
52 PRIOR APPLICATION NUMBER: 60/083559
53 PRIOR FILING DATE: 1998-04-29
54 PRIOR APPLICATION NUMBER: 60/083500
55 PRIOR FILING DATE: 1998-04-29
56 PRIOR APPLICATION NUMBER: 60/083742
57 PRIOR FILING DATE: 1998-04-30
58 PRIOR APPLICATION NUMBER: 60/084366
59 PRIOR FILING DATE: 1998-05-05
60 PRIOR APPLICATION NUMBER: 60/084414
61 PRIOR FILING DATE: 1998-05-06
62 PRIOR APPLICATION NUMBER: 60/084441
63 PRIOR FILING DATE: 1998-05-06
64 PRIOR APPLICATION NUMBER: 60/084637
65 PRIOR FILING DATE: 1998-05-07
66 PRIOR APPLICATION NUMBER: 60/084639
67 PRIOR FILING DATE: 1998-05-07
68 PRIOR APPLICATION NUMBER: 60/084640
69 PRIOR FILING DATE: 1998-05-07
70 PRIOR APPLICATION NUMBER: 60/084598
71 PRIOR FILING DATE: 1998-05-07
72 PRIOR APPLICATION NUMBER: 60/084600
73 PRIOR FILING DATE: 1998-5-07

Query Match	36.38;	Score 808;	DB 9;	Length 473;
Best Local similarity	43.38;	Score 808;	DB 9;	Length 473;

conduct vacuoles 30; mismatches 140; indels 84; gaps 12;

[illegible]

77 PGFGS--NLLTLWLFSSNNLSTIYPGTFRHLQALEELDLCGDNBRHLBSLEPDTFOGI.FRI.O 13

OV 135 SL.H.YRCOI.SSI.BENTFBJVSLQVI.VI.QENST.I.W.OOOF.EBET.NW.AWA.....

Db 134 TLHLDRCGLQELGPGLE*RGLAALQYLYLQDNALQALPDDTFERDLGNLTHLFLHGNRISSV 19

Db 194 PERAFGLSLDRLLHONRVAAVHPHAFBDIGRI.MTYIFANNI SAI PTEAT ADI PATC CC

255 FLRLNANPWACDCRARPLWAFQRRARVSSSDVTCA TPPEGRDLRALREADFQAC --- 31

[illegible]

514 GPTHTWIGRAIDEEPLGLPKCCQPDADDKASVLEPGRPASAGNALKGRVPPGDSPPGNG 37

374 — — — — —

QY 362 GDAPTEDDYWG YGGEDQ RGEQMCPGAACQAPP 394

Sequence 382, Application US/10028072

APPLICANT: Baker, Ke

APPLICANT: Beresini, Maureen

APPLICANT: Desnoyers, Luc
APPLICANT: F41-2004-011

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Mei-Qiang
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gunney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang
 TITLE OF INVENTION:
 FILE REFERENCE:
 CURRENT APPLICATION NUMBER: US/10/028, 072
 PRIOR FILING DATE: 2001-12-19
 PRIOR APPLICATION NUMBER: 60/049911
 PRIOR FILING DATE: 1997-06-18
 PRIOR APPLICATION NUMBER: 60/056974
 PRIOR FILING DATE: 1997-08-26
 PRIOR APPLICATION NUMBER: 60/059113
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/059115
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/059117
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/059122
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/059184
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/059263
 PRIOR FILING DATE: 1997-09-18
 PRIOR APPLICATION NUMBER: 60/059352
 PRIOR FILING DATE: 1997-09-19
 PRIOR APPLICATION NUMBER: 60/059588
 PRIOR FILING DATE: 1997-09-19
 PRIOR APPLICATION NUMBER: 60/059836
 PRIOR FILING DATE: 1997-09-24
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/062285
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/062287
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/062814
 PRIOR FILING DATE: 1997-10-24
 PRIOR APPLICATION NUMBER: 60/062816
 PRIOR FILING DATE: 1997-10-24
 PRIOR APPLICATION NUMBER: 60/063045
 PRIOR FILING DATE: 1997-10-24
 PRIOR APPLICATION NUMBER: 60/063082
 PRIOR FILING DATE: 1997-10-31
 PRIOR APPLICATION NUMBER: 60/063127
 PRIOR FILING DATE: 1997-10-24
 PRIOR APPLICATION NUMBER: 60/063327
 PRIOR FILING DATE: 1997-10-27
 PRIOR APPLICATION NUMBER: 60/063329
 PRIOR FILING DATE: 1997-10-27
 PRIOR APPLICATION NUMBER: 60/063550
 PRIOR FILING DATE: 1997-10-28
 PRIOR APPLICATION NUMBER: 60/063561
 PRIOR FILING DATE: 1997-10-28
 PRIOR APPLICATION NUMBER: 60/063704
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063733
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063735
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063738
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063755
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/064248

PRIOR FILING DATE: 1997-11-03
 PRIOR APPLICATION NUMBER: 60/064809
 PRIOR FILING DATE: 1997-11-07
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065846
 PRIOR FILING DATE: 1997-11-17
 PRIOR APPLICATION NUMBER: 60/066364
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: 60/066453
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/066511
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/069212
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069278
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069334
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069694
 PRIOR FILING DATE: 1997-12-16
 PRIOR APPLICATION NUMBER: 60/072320
 PRIOR FILING DATE: 1998-01-23
 PRIOR APPLICATION NUMBER: 60/073612
 PRIOR FILING DATE: 1998-02-04
 PRIOR APPLICATION NUMBER: 60/074086
 PRIOR FILING DATE: 1998-02-09
 PRIOR APPLICATION NUMBER: 60/074092
 PRIOR FILING DATE: 1998-02-09
 PRIOR APPLICATION NUMBER: 60/077791
 PRIOR FILING DATE: 1998-03-12
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079663
 PRIOR FILING DATE: 1998-02-27
 PRIOR APPLICATION NUMBER: 60/079728
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081695
 PRIOR FILING DATE: 1998-04-14
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081818
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082999
 PRIOR FILING DATE: 1998-04-24
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085149
 PRIOR FILING DATE: 1998-05-12
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13


```

: PRIOR APPLICATION NUMBER: 60/085579
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085697
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085704
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/086414
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/086430
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087106
: PRIOR FILING DATE: 1998-05-28
: PRIOR APPLICATION NUMBER: 60/088026
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088730
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088741
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088810
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088858
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/089532
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089599
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089907
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089947
: PRIOR FILING DATE: 1998-06-19
: PRIOR APPLICATION NUMBER: 60/090349
: PRIOR FILING DATE: 1998-06-23
: PRIOR APPLICATION NUMBER: 60/090429
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090445
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090338
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090863
: PRIOR FILING DATE: 1998-06-26
: PRIOR APPLICATION NUMBER: 60/091360
: PRIOR FILING DATE: 1998-07-01
: PRIOR APPLICATION NUMBER: 60/091519
: PRIOR FILING DATE: 1998-07-02
: PRIOR APPLICATION NUMBER: 60/091982
: PRIOR FILING DATE: 1998-07-07

```

Query Match 36.3%; Score 808; DB 9; Length 473;
 Best Local Similarity 42.2%; Pred. No. 1,1e-57;

Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

```

OY 18 LMLALPLAAPSCLMCTCYSSP-TVSCQANNSSVPLSLPSTORLFLONNLIRTLR 76
DB 15 VLMLOAMQVAAP-CPGACVCYNKPKVTTSCTPOGQLOAVPGVIGIPASORIFLHGNRISHVP 73
OY 77 PGTGSS--NLTLMLFSSNNLSTIYPGTFRHLOALEEIDLGNRHLRSLEPDTFOGLERLO 134
DB 74 AASFRACRNLTLMLHSHVNLARIDAAFTGLALEOGLDLSNQAOLRSVDPATFPGHGLRLH 133
OY 135 SLHLRYCOLSLSPGNIFRGVLSLOYLYLOENSLLHLODDLFAFLANLSHLFLHGNRLRL 194
DB 134 TLHLDRCGLOELRGFLGLAOLYLQLDNALQALPDDTFRRDLGNLTHLFLHGNRISYSV 193
OY 195 TEHVFRGLSLDRLLHGNRLQGVHRAAFRGLSRLLTYLFLFNNSLASLPGCALADLSLE 254
DB 194 PERAFRGHSLDRLLHGNRVAVHHPAFRDLGRMLTYLFAANNLSALPTALALPLRALQ 253
OY 255 FLRLNANWACDCRRARPLAMFQBARVSSDYTCATPPEPGRDRLRALREADFOAC---- 310
DB 254 YLRINDPWCDCRRARPLAMWLQKFRGSSSEVPCLPQRLAGRLKRLAANDLOGCAVAT 313
OY 311 -----PPAA-----PTRGSRAR-----GNS 326

```

```

DB 314 GPHYPIWTRATDEEPLGLPKCCORDPAADKASYLEBGRPASGNMLKGRVPPGDSPPGNG 373
OY 327 SS-NHL-----YGAAGAPADPSTLYR-----DLPAEDSRGRG----- 361
DB 374 SGRRHINDSPFGLTLPGSABP--PLTAVRPEGSEPPGFPPTSGRRRRRRCSSKRRRTSHOCL 431
OY 362 GAAPTEDDYWGXYGGEDQGEQMCPCGACQARP 394
DB 432 GQAGS-----GGGCTGDESGALPLSLTCSLNP 459

```

RESULT 7
 US-10-121-049-382

; Sequence 382; Application US/10121049
 ; Publication No. US2003002239A1

; GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tunas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P330R1C17
: CURRENT APPLICATION NUMBER: US/10/121,049
: CURRENT FILING DATE: 2002-04-12
: PRIOR APPLICATION removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 382
: LENGTH: 473
: TYPE: PRT
: ORGANISM: Homo Sapien
: US-10-121-049-382

```

Query Match 36.3%; Score 808; DB 9; Length 473;
 Best Local Similarity 42.2%; Pred. No. 1,1e-57;

Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

```

OY 18 LMLALPLAAPSCLMCTCYSSP-TVSCQANNSSVPLSLPSTORLFLONNLIRTLR 76
DB 15 VLMLOAMQVAAP-CPGACVCYNKPKVTTSCTPOGQLOAVPGVIGIPASORIFLHGNRISHVP 73
OY 77 PGTGSS--NLTLMLFSSNNLSTIYPGTFRHLOALEEIDLGNRHLRSLEPDTFOGLERLO 134
DB 74 AASFRACRNLTLMLHSHVNLARIDAAFTGLALEOGLDLSNQAOLRSVDPATFPGHGLRLH 133
OY 135 SLHLRYCOLSLSPGNIFRGVLSLOYLYLOENSLLHLODDLFAFLANLSHLFLHGNRLRL 194
DB 134 TLHLDRCGLOELRGFLGLAOLYLQLDNALQALPDDTFRRDLGNLTHLFLHGNRISYSV 193
OY 195 TEHVFRGLSLDRLLHGNRLQGVHRAAFRGLSRLLTYLFLFNNSLASLPGCALADLSLE 254
DB 194 PERAFRGHSLDRLLHGNRVAVHHPAFRDLGRMLTYLFAANNLSALPTALALPLRALQ 253
OY 255 FLRLNANWACDCRRARPLAMFQBARVSSDYTCATPPEPGRDRLRALREADFOAC---- 310
DB 254 YLRINDPWCDCRRARPLAMWLQKFRGSSSEVPCLPQRLAGRLKRLAANDLOGCAVAT 313
OY 311 -----PPAA-----PTRGSRAR-----GNS 326
DB 314 GPHYPIWTRATDEEPLGLPKCCORDPAADKASYLEBGRPASGNMLKGRVPPGDSPPGNG 373

```

OY 327 SS-NHL-----YGAAGAPADPSTLYR-----DLPAEDSRGROG----- 361
 Db 374 SGRHINDSPFGITLPGSAEP--PLTAVRPEGSEPPGPTSGPRRRRPGCSRKNNRTRSHCRL 431
 OY 362 GDAPTEDDYWGCGEGEDQGEOMCPGAACQAP 394
 Db 432 GQAGS-----GGGTGDESGGALPILTCSLTP 459

RESULT 8
 US-10-123-904-382
 ; Sequence 382, Application US/10123904
 ; Publication No. US20030022328A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowsky, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3330R1C34
 CURRENT FILING DATE: 2002-04-16
 PRIOR APPLICATION REMOVED - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 382
 LENGTH: 473
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-123-904-382

Query Match 36.3% Score 808; DB 9; Length 473;
 Best Local Similarity 42.2% Pred. No. 1,1e-57;
 Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

18 LMLTALPLAAPSCLCTCYSSP-TVSCQANNFSSVPLSLPSTQRLFLQNNLITRLR 76
 Db 15 VMLQAMQVAAP-CFACVCYNEPKVTTCPOGQAVPVGIPASQRIFLHGNRISHVP 73
 OY 77 PGTFGS--NLTTWLFSSNNLSTYPGTFRLQALDELIDGNRHLRSLDEPTFOGLERLQ 134
 Db 74 AASFRCRNLTITLHNSVNLARIDAAAFGLALDELIDSDNOLSVDPATFHGRLH 133
 OY 135 SLHLYRCQSLPLPGNIFRGVLSQYLYLOENSLHLHODDLFADLANLSHLFLHGNRLRL 194
 Db 134 TLHLDRCGLQELGFLRGIAALQYLYLODNALQALPDDTFRDLGNLTHFLHGNRISSV 193
 OY 195 TEHVFRGLSLDRLLHGNRLQGVHRAAFRGLSRLTLYLFNNSLASLDEALADLPSE 254
 Db 194 PERAFRGLSHLDRLLHGNRVAVHRAFRDLGRMTLYLFANNLSALPTEALAPRALQ 253
 OY 255 FLRLANPMAQCDRARPLWAMFORAVSSSDVTCATPPERQGRDLRLREADFOAC---- 310
 Db 254 YLRINDPWCDCRARPLWAMLOKFRGSSSEVPQSLPQRLAGRLKRLAANDLQGCATVAT 313
 OY 311 -----PPAA-----PTRGSSRAR-----GNS 326
 Db 314 GPHYHPTWGRATDEPLGLPKCCQPDAAADKASVLEGRPASAGNALKGRVPGDSPGNG 373

OY 327 SS-NHL-----YGAAGAPADPSTLYR-----DLPAEDSRGROG----- 361
 Db 374 SGRHINDSPFGITLPGSAEP--PLTAVRPEGSEPPGPTSGPRRRRPGCSRKNNRTRSHCRL 431
 OY 362 GDAPTEDDYWGCGEGEDQGEOMCPGAACQAP 394
 Db 432 GQAGS-----GGGTGDESGGALPILTCSLTP 459

RESULT 9
 US-10-140-470-382
 ; Sequence 382, Application US/10140470
 ; Publication No. US20030022331A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowsky, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3330R1C160
 CURRENT FILING DATE: 2002-05-06
 PRIOR APPLICATION REMOVED - See Palm or File Wrapper
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 382
 LENGTH: 473
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-140-470-382

Query Match 36.3% Score 808; DB 9; Length 473;
 Best Local Similarity 42.2% Pred. No. 1,1e-57;
 Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

18 LMLTALPLAAPSCLCTCYSSP-TVSCQANNFSSVPLSLPSTQRLFLQNNLITRLR 76
 Db 15 VMLQAMQVAAP-CFACVCYNEPKVTTCPOGQAVPVGIPASQRIFLHGNRISHVP 73
 OY 77 PGTFGS--NLTTWLFSSNNLSTYPGTFRLQALDELIDGNRHLRSLDEPTFOGLERLQ 134
 Db 74 AASFRCRNLTITLHNSVNLARIDAAAFGLALDELIDSDNOLSVDPATFHGRLH 133
 OY 135 SLHLYRCQSLPLPGNIFRGVLSQYLYLOENSLHLHODDLFADLANLSHLFLHGNRLRL 194
 Db 134 TLHLDRCGLQELGFLRGIAALQYLYLODNALQALPDDTFRDLGNLTHFLHGNRISSV 193
 OY 195 TEHVFRGLSLDRLLHGNRLQGVHRAAFRGLSRLTLYLFNNSLASLDEALADLPSE 254
 Db 194 PERAFRGLSHLDRLLHGNRVAVHRAFRDLGRMTLYLFANNLSALPTEALAPRALQ 253
 OY 255 FLRLANPMAQCDRARPLWAMFORAVSSSDVTCATPPERQGRDLRLREADFOAC---- 310
 Db 254 YLRINDPWCDCRARPLWAMLOKFRGSSSEVPQSLPQRLAGRLKRLAANDLQGCATVAT 313
 OY 311 -----PPAA-----PTRGSSRAR-----GNS 326
 Db 314 GPHYHPTWGRATDEPLGLPKCCQPDAAADKASVLEGRPASAGNALKGRVPGDSPGNG 373
 OY 327 SS-NHL-----YGAAGAPADPSTLYR-----DLPAEDSRGROG----- 361

Db 374 SCPRINSPFGTLPGSAEP--PLTAVRPESEEPGPFPTSGPRRRPGCSKKNRTSRCL 431

QY 362 GDAPTEDDYWGCGEGEDQRGFOMPGAACQAP 394

Db 432 GQAGS-----GGCGTGDESGGALPSLTCLTP 459

RESULT 10
US-10-175-746-382

; Sequence 382, Application US/10175746
; Publication No. US20030027270A1
; Publication No. US20030027270A1

```

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P93330RIC353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 382
LENGTH: 473
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-746-382

```

Query Match 36.3%; Score 808; DB 9; Length 473;
 Best Local Similarity 42.2%; Pred. NO. 1.1e-57;
 Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12

0Y	18	LLALALPLAASCPMYLCTCTSSP--TVSCQANFNSSVPLSLRPSNORLEFLONNLIRTLR	76
0Y	15	VLMIAQNOVAAP--CPGACVCVTEPKYKTTSCPPQOGLQAPVIGTAAISORITPLHGNSIHPV	73
0Y	77	PGTFGS--NLTLMLSPNSNLSTVYPGTFPHLQALEETLDLDNHRKLSLEDPFGGLEHQ	134
Db	74	AASEFACRNLTTITLHNSNYLARIDAFAFTGLALLEQDLSDNMQILRSVPATPHGIGRLH	133
0Y	135	SLHLHYRCQSLSSLEGNFERGIVSYLOLYLXLDQNSLLHLODDLFADLANLSHFLHGNLRLL	194
Db	134	TLHLDRCGLOELGPGLERGALQALQYLYLDONALQALRPDTFRQLGMLTFLHGNRISSV	193
0Y	195	TEHVFRGSLDRLLLHGNRLQGVHNAAFRGSLRTILYLFNNLSLSPGEALADPSE	254
Db	194	PERAFRGSLHSLDRLLIQNVAVVHHPAFRDLGRMLTYLYFANNLSALPTALAPRLAQ	253
0Y	255	FLRLNANPMACDCRARPLAMFGRARVSSSDVYCAITPRPEQCGDLDALKEADPQAC---	310
Db	254	YLRINDPMVCCDCRARPLAMLQKFKGSSEVPCSLPRLAGLADIKRLANDLQCAVAT	313
0Y	311	-----PPAA-----PTRGSGRAR-----	326
Db	314	GPYHPITWGRATDEEPLGLPKCCQPPAADKASVLEBPGRPASAGNALKGVPEDSPGNG	373
0Y	327	SS-NHL-----YGVAAEGARPAPDSTLYR-----DLPAEDSGRQG-----	361

Db 374 SGPRINDSPGTLPGAEP---PLTAVRPEGSEPPGPTSGPRRRPGCSRKNRTSRHRL 431

OY 362 GDAPTEDDYMGVGGEDQRGEQMGGAACQAPP 394

Db 432 GQAGS-----GGGGGCDSESGALPSLTCSLTP 459

RESULT 11
US-10-176-918-382

; Sequence 382, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:

```

1  GENERATED INFORMATION:
2  APPLICANT: Baker, Kevin P.
3  APPLICANT: Beresini, Maureen
4  APPLICANT: DeForge, Laura
5  APPLICANT: Desnoyers, Luc
6  APPLICANT: Filvaroff, Ellen
7  APPLICANT: Gao, Wei-Qiang
8  APPLICANT: Gerritsen, Mary E.
9  APPLICANT: Goddard, Audrey
10 APPLICANT: Godowski, Paul J.
11 APPLICANT: Gurney, Austin L.
12 APPLICANT: Sherwood, Steven
13 APPLICANT: Smith, Victoria
14 APPLICANT: Stewart, Timothy A.
15 APPLICANT: Tumas, Daniel
16 APPLICANT: Watanabe, Collin K
17 APPLICANT: Wood, William
18 APPLICANT: Zhang, Zemin
19
20 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
21 FILE REFERENCE: P3330R1C382
22 CURRENT APPLICATION NUMBER: US/10/176,918
23 CURRENT FILING DATE: 2002-06-20
24
25 Prior Application removed or Palm
26 NUMBER OF SEQ ID NOS: 550
27
28 SEQ ID NO 382
29 LENGTH: 473
30
31 TYPE: PRF
32
33 ORGANISM: Homo Sapien
34
35 US-10-176-918-382

```

Query Match	36.3%	Score 808;	DB 9;	Length 473;
Best Local Similarity	42.2%	Pred. No. 1.1e-57;		
Matches 191; Conservative	38;	Mismatches 140;	Indels 84;	Gaps 12;

QY	18	LIMLALPLAASCPMYCTCTGSSP--TVSCQANFNSSVPLSLPSTPQRLFLQNLNLTIRL	76
Db	15	VLMIQANQVAAP--CPGACVCTNEPKYTTSPDQGLAAYVPGIAPASQRIPLHGNRISHP	73
QY	77	PGTFGS--NLTLMLFNSNNLSTTPYPGTFPHQLQALEEDLQDNRLSLSPDFQGLERQ	134
Db	74	AASPRACRNLTILMLHSNVLARIDAAFTGLALBEDQLSDMQLSVDPATPHGICRLH	133
QY	135	SLHLHYRQQLSSPEQNTFRGLSVLSQVLYLQDENSLLHQQDLFADLANLSHLFLHGRLLL	194
Db	134	TLHLDRGQLELQGLERFGLEAALQYLYLQNALQALPDDTFRDLGMLTHLFLHGRISSV	193
QY	195	TEHYFRGLSGIDRLLLHGNGRLQGVNHAAPFGISRLITVLFNNSLASLQGEALADLPSE	254
Db	194	PERAFRGSLHDLRLTLHQNVAHVNHAFEDLGRLLMTVLYLFANLSALPTLEALPLRALQ	253
QY	255	FLRLNAPMAACDCRARPLTMAFQARVASSDQVCAATPERQGRDLALREADFOAC----	310
Db	254	YLRINDPVMVCCDCRARPLTMAWLQKPFSSSEVFCSLPQRLAGIDLKRLAANDLQCAVAT	313
QY	311	-----PRAA-----PTRGGRAR-----GNS	326
Db	314	GYHHPIWTGRATDEEPLGLPKCCQPPAADKASVLEBGRASAGNALKGAVPCDSPGNG	373
QY	327	SS--NHL-----YGVAAEGARPADPSTLYR-----DLPAEDSGRGQ-----	361
Db	374	SGRHHINDSFGLPGSADE--PLTVAREPGSEPGCFPTSGPRRRQCSGRKKNPTNSHCL	431


```

PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23

```

```

PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835

```

Query Match 36.3%, Score 808, DB 9, Length 473;
 Best Local Similarity 42.2%, Pred. No. 1, le-57;
 Matches 191; Conservative 38; Mismatches 140; Indels 84; Gaps 12;

```

QY 18 LMLLALPLAASCPMLTCYSSP-TVSCQANNSSVPLSLPSTQRLPLQNNLIFTLR 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db- 15 VMLAMOVAAAP-CPGACVCINERKVTTSQPOQGLQAVPIPAASORIFLHGMRISHVP 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 PGTGGS--NLTLWLESNNLSTIYPGTFRHQLQALELDLGDNRHLSLEPTFOGLERLQ 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 AASFRCRNLTLTLMHSHVRLDAAFDTGLALLEQDLSDNADLRSDVDPATFHLGRLLH 133

```



```

Db 134 TLHIDRCGLQELGGLFRGLAALQYLYLDONALQALPDDTFRDGLNLTFLHGNRISSV 193
OY 195 TEHVFRGLGSLDRLLHGNHRAAFRGSLRTILYLFNNSLASLPGALADLPSTL 254
Db 194 PERAFRGLHSLDRLLHGNHRAAFRGSLRTILYLFNNSLASLPGALADLPSTL 253
OY 255 FLRLNANPMACDCRANPLMAMFORARVSSDYTCATPEROGRDLRALREADFOAC 310
Db 254 YLRINDNPMWCDRCRANPLMAMLOKFRGSSSEVPCLPQRLAGRDILKRLANDLOGCAVAT 313
OY 311 -----PPAA-----PTRGSRAR-----GNS 326
Db 314 GRYPRTMTGRATDEERLGLPKCCQPDADKASVLEPRGASAGNALKGRVPPGDSPPGNG 373
OY 327 SS-NHL-----YGAERAGAPRADPSTLYR-----DLPAEDSRGROG----- 361
Db 374 SGRRHINDSPFGTLPGSAEP--PLTAVRPEGESEPRGFTSGRRRPGCSRKNRTRSHCRL 431
362 GDAPTEDDYWGYYGGEDQGEQMCPCAACQAP 394
432 GQAGS-----GGGTGDSBEGSALPSTLCSLTP 459

```

Search completed: April 4, 2003, 12:11:50
 Job time : 22 secs


```

OY 84 -----LLTLMFSSNNLSTIYPGT 101
Db 60 LDTQLCFRPPDARFGLPRLLEDLEVTGSSFLNLTSTSLGKLTLPNFMLEALPEGL 119
OY 102 FRHLQALEELDLGDNHRLSLPDTFOGLERLQSLHLYRCQLSSLPGNIFRGIVSLQYLY 161
Db 120 FOHLAALLESILQGN-LQALPRRLFOPLTHLKTNLQNLQAOLPELPHPLTSLQTLK 178
OY 162 LOENSLHLQDDLFADLANLSHLFLHGNRLRLTEHVFRGLSLDRLLHGNRLQGVHRA 221
Db 179 LSNNAISGLPQGVFGKLSLQELFLDSNNISSELPQVFSOLFCLERLWLQRNAITHPLS 238
OY 222 AFRGLSRITILYLFNNSLASLPGCALADLPSEELRLMAN 261
Db 239 IFASLGNLTFTLSLQNMMLRVLPAGLFAHTPCIVGLSLTHN 278

```

Search completed: April 4, 2003, 12:10:50
 Job time : 50 secs

F:202-225/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>
 F:226-249/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>
 F:262-309/Domain: proteoglycan carboxyl-terminal homology <PCH>
 F:2/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F:8-21,268-294/Disulfide bonds: #status experimental
 F:44,151,234,290/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:271/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 12.6%; Score 280.5; DB 1; Length 312;
 Best Local Similarity 30.6%; Pred. No. 1.8e-13;
 Matches 95; Conservative 41; Mismatches 135; Indels 39; Gaps 8;

32 PMLCTCYSSP--PTVSCQANNFSSVPLSPST-----ORL 65
 5 PKDCOVFSDHSSISQCP--AEIPGYLPATVHLAVEFNLTHPLNLOGASKDEL 62
 DB
 66 FLQNNLIRTLRPGTEG--SNLTLWLFSSNNLSTIYPGTFRHLOALEELDGNRLRSIE 123
 DB
 63 HLSSNGLESLEPEFLRPVQLRVLDTRNALGLPPGLEFOASATLDTLVKENO--LEVLE 121
 124 PPTFGLERLQSLHYRCOLSLPGNIFRGVLSLOYLYIQENSLSLHLODDLFAADLANLSH 183
 DB
 122 VSMILGKRALGHLDSGNRLRPLPGLANFTLRLDGENQLETPPLDLRGPLQDER 181
 184 LFLHGNRLRLTEHYFRGSLDRLLHGNRLQGVHRAAFRGLSRLTLLEFNNSLASLP 243
 DB
 182 LHLBNKRLQVGLKDLLPQDLRYLFLNKNKLARAAGFQGLRDLMDLSDNNSLASVP 241
 QY
 244 GHALADL--PSLEF--LRNANPNACDCRARPANAFORAR---VSSSDVTCATPPEQ 295
 DB
 242 EGIAMSLGPNMDKRGDFISGNPMICDNLSDLYKMLQAKDKMFSQNDTRCAGPEAVK 301
 QY
 296 GRDLRALREA 305
 DB
 302 GQTLAVAKS 311

RESULT 14

A58532
 gital cell membrane glycoprotein LG-1 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999
 C:Accession: A58532
 R:Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.
 J. Biol. Chem. 271, 22522-22527, 1996
 A:Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in A:Reference number: A58532; MUID:96394313; PMID:8798419
 Accession: A58532
 Status: Preliminary; translated from GB/EMBL/DBJ
 Molecule type: mRNA
 A:Residues: 1-1091 <SUZ>

A:Cross-references: GB:D78572; NID:91545806; PIDN:BA11416.1; PID:91545807
 C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-ter
 F:36-61/Domain: leucine-rich alpha-2-glycoprotein repeat homology <PCH>
 F:71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:166-188/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 12.5%; Score 278.5; DB 2; Length 1091;
 Best Local Similarity 27.0%; Pred. No. 1.1e-12;

Matches 99; Conservative 53; Mismatches 126; Indels 89; Gaps 10;

38 YSSPTVSCQANNFSSVPLSPSTQR---LFLQNNLIRTLRPGTE---GSNLTLMFLS 91
 DB
 140 YLSLEVLDSNNITTEINSSCPNGLRRELNLASNRISLESAGFDLSRLRLRSK 199
 QY
 92 NNLSTIYPGTFRHLOALEELDGNRLRSLEPPDQGLERLQSLHYRCOLSLPGNIF 151
 DB
 200 NRIITLQPKAFK-LPRLQDLNRRN-IRLIGLTFQGLDSEVLRLQRRNNISRLTDAF 257
 QY
 152 RGLVSLQYLYIQENL-----LH----- 169
 DB
 258 WLSKMHVLLHLYNSLVEVNSGLYGLTALHQLHLSNNSISRIQDQGSFCKLHELIS 317
 QY
 170 -----LQDDLFAADLANLSHLFLHGNRLRLTEHYFRGSGS----- 204
 DB
 318 FNNLIRLDEESLAEISSILRLSHNAISHLBGAFKLSRLVLDLHNEISGTEEDTS 377
 QY
 205 -----LDRLLHGNRLQGVHRAAFRGLSRLTLLEFNNSLASLPGEALDPSLEFL 256
 DB
 378 GAFGLDNLSTKLTLFGNKLKSVAKRAFSGLSELEHNLGEMAIRSVQPAFAKMKNLKEL 437
 QY
 257 RLANPNACDCRARPANAFQ--RAVSSSDVTCATPPEQGRDLRALRADPQAC---P 311
 DB
 438 YISSSEFLDCQIKLPLWLMGRMLQAEVYATCAHPESLKGSTFSLPDSF--VCDDEPK 496
 QY
 312 PAAPTRP 318
 DB
 497 PQTITQP 503

RESULT 15

A34901
 lysine carboxypeptidase (EC 3.4.17.3) 83K chain - human
 C:Species: Homo sapiens (man)
 C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 05-Nov-1999
 C:Accession: A34901
 R:Tan, F.; Weerasinghe, D.K.; Skidgel, R.A.; Tamei, H.; Kaul, R.K.; Roninson, I.B.; S
 J. Biol. Chem. 265, 13-19, 1990
 A:Title: The deduced protein sequence of the human carboxypeptidase N high molecular
 A:Reference number: A34901; MUID:9009486; PMID:2378615

A:Residues: 1-536 <TAN>
 A:Cross-references: GB:J05158; NID:9179935; PIDN:AAA51921.1; PID:9179936
 Accession: A34901
 Status: preliminary
 A:Molecule type: mRNA
 A:Gene: GDB:ACBP
 A:Cross-references: GDB:127893
 A:Map position: 6q25.3-6q26
 C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
 C:Keywords: hydrolase; metallo-carboxypeptidase
 F:77-100/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:221-244/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:245-268/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:269-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:293-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:317-340/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:341-364/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>

Query Match 12.3%; Score 273.5; DB 2; Length 536;

Best Local Similarity 30.7%; Pred. No. 1.1e-12;

Matches 86; Conservative 25; Mismatches 116; Indels 53; Gaps 3;

32 PMLCTCYSSPTVSCQANNFSSVPLSPSTQRFLQNNLIRTLRPGTEGSSN----- 83
 DB
 2 PKGCDCEVQ--EYFCSDDELATVPLDIPYTKNIIVERSTLTLEFRAGSNPNLKVVE 59

QY 310 CPPAATPFGSRRARGNSSNNHYGVAEAG--APPAD-PSLT 347
 Db 271 QVPACTLSSGSCFAMCSCNGIYDCRGKGLTAIPALPFTM 311

RESULT 11

T13953

MEGF5 protein - rat

N:Alternate names: slit protein homolog

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002

C:Accession: T13953

R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.

A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs

A:Reference number: 24126; MUID:98360089; PMID:9693030

C:Accession: T13953

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1523 <NAK>

A:References: EMBL:AB011531; NID:g3449291; PIDN:BA032461.1; PID:g3449292

A:Gene: MEGF5

C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

Query Match 13.2% Score 293; DB 2; Length 1523;

Best Local Similarity 21.2%; Pred. No. 1.4e-13;

Matches 103; Conservative 59; Mismatches 144; Indels 180; Gaps 10;

QY 1 MPEGLRRLQAPASACLLMLLPLAAP---SCPMLCTCYSSPPTVSCQANNFSSVPLS 57

Db 1 MAPRGAGAAVARRALALALASISGPPAACPTKCTC--SASVSDCHGGLRAVPNG 58

QY 58 LPPSTQRLFL-QNNLIRLTPGTFG-SNLTLTLESNNLSTIYPTGFRHLQALELDLG- 114

Db 59 IPRNAERLDLDNNITRTIKMDFTGLKNRLVLEEDNQSVIERGAFQDLKQLERLRLNK 118

QY 115 -----DNHLSLSPEDTFQ 128

Db 119 NKLQVLPPELLFQSTPKLTLDLSENQIOGIPKKAERGVYGVKNLQDLDNHNISCIEDGAGR 178

QY 129 GLE-----RLQSLHL----- 139

Db 179 ALRDLLEILLNNNISRLIVTSPNHPKIRTLRLHNSHLKYCHLAWLMDLRLQRTIQ 238

QY 140 -----RCQLSSL----- 146

Db 239 FTLCMAPVHLRGFSVADYQKKEVYCPGPHSEAPACNANSLSCPSACSCNNIYDCRGKL 298

QY 147 ---PGNIFGLVS-----LOY-----LYLQENSLLHODDLFADLANLS 182

Db 299 TEIPANLRPGIYEIRLEONSISIPAGAFIYKKLKRIDISKNOISDIAPDAFQGLSKLT 358

QY 183 HFLHGNRLRLTEHFVFRGLSLDRLLHGNRLQGVHRAFGRLSLITLYLFNNSLASL 242

Db 359 SLIVGNKITEIRPKGLFDGLVSLQILLNLANKINCLRVATFQDQLNLNLSLYDKMLQTI 418

QY 243 PEEALADPLSEFLRLNANPACDCRARPILAMFORARVSSSDVYCATPEPQGRDLRL 302

Db 419 SKGLVAPLQSIQTLHLAQNPFVCDLKLWLDYLDNPIETSGARCSSPRRLANKRISOI 478

QY 303 READPQ 308

Db 479 KSKKFR 484

Db 479 KSKKFR 484

RESULT 12

JC7763

neural leucine-rich repeat protein-3 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: JC7763

R:Fukumachi, K.; Matsuo, Y.; Kitahara, C.; Kuchino, Y.; Tsuda, H.

Biochem. Biophys. Res. Commun. 287, 257-263, 2001

A:Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the ge

A:Reference number: JC7763; PMID:11549284

A:Contents: Fibrosarcoma cells

A:Accession: JC7763

A:Molecule type: mRNA

A:Residues: 1-707 <FUK>

A:Cross-References: CR:AF291437

C:Comment: This protein, a new member of the neuronal leucine-rich repeat protein fam

C:Genetics:

A:Gene: nlrp-3

C:Keywords: cell adhesion

Query Match 12.9% Score 286.5; DB 2; Length 707;

Best Local Similarity 22.8%; Pred. No. 1.7e-13;

Matches 108; Conservative 53; Mismatches 165; Indels 147; Gaps 10;

QY 20 MLLALPLAA-----PSCPMLCTC-----YSSPPTVSCQANNFSSVPLSP 60

Db 10 VLLGLAIFALVQAGDKKVDPCQLCTCEIRPWFTPRSIYEAETVDCNDGLNLPARLPA 69

QY 61 SNORLEFLQ-NNLIRLTPGTFGSLNLTLMFESNNLSTIYPTGFRHLQALELDLQDN-- 117

Db 70 DTQILLQTNNTARLEHSTDFPVNLTGLDLSQNNLSVYINIVQKMSQLSVYLEENKLT 129

QY 118 -----H-LSLEPPTFGLERLQSLHLYRCQLSLPGNIFGLVS 156

Db 130 ELPEKCLVGLSNLQELVYVHNLNLSAISPGAFVGLNLRHLHNSRLQDINKWEALPRL 189

QY 157 LQYLYQENSLLHODDLFADLANLSHLFAGNRRLRLTEHFVFRGLGSLDRLLHGNRLQ 216

Db 190 LEILMGNDPILIRKIDMNFQPLKRLSLVIAGINTEVPDDALVGLLENLSTISFYDNRLN 249

QY 217 GV----- 218

Db 250 KVPYALQKAVNLKFLDLNKNPINKIRGDFSNMLKLELGINNPELYSIDSLAVNDLP 309

QY 219 -----HRAFRGLSLRLTYLFNNSLASLPGALADPLSEFLRLNMP 262

Db 310 DLKITEATNPRLSYIHPNAFRLPKLESLMNSALSLVYHGTIESLPNKEISHSNP 369

QY 263 WACDCRARPILAMFORARV---SSSDVYCATPEPQGRDLRLREAD-FQACPPAAPTRP 318

Db 370 IRDCDVIR--WINMKNTIRFMEPDSLFCVDPPEQGVQVVRVHFMDEICLPLT----- 423

QY 319 GSRAGNSSNNHYGVAEAGAPPADPSTLYRDLPAEDSRGCGDAPTEDDYW 371

Db 424 -----APESFPIIDVADSVSLHCNATAPQPEIYW 456

RESULT 13

NBH0A2

leucine-rich alpha-2-glycoprotein - human

C:Species: Homo sapiens (man)

C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 05-Dec-1998

C:Accession: A03211

R:Takahashi, Y.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 82, 1906-1910, 1985

A:Title: Periodicity of leucine and tandem repetition of a 24-amino acid segment in c

A:Reference number: A03211; MUID:85166241; PMID:3858686

A:Accession: A03211

A:Molecule type: protein

A:Residues: 1-312 <TAK>

C:Comment: The function of this plasma protein is not known.

C:Superfamily: leucine-rich alpha-2-glycoprotein; leucine-rich alpha-2-glycoprotein r

C:Keywords: duplication; glycoprotein; plasma; tandem repeat

F:58-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:106-129/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:130-153/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:154-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:178-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

Db 301 NNPKKCTCQLRGLRRLWE-AKTSRPDATCASPARKRGH---IRDTDAFRGC--KFFIKR 354

QY 319 GSRA 322
Db 355 SKKA 358

RESULT 9

A60164
platelet membrane glycoprotein V precursor - human
C:Species: Homo sapiens (man)
C>Date: 12-Jan-1993 #sequence_revision 24-Feb-1994 #text_change 05-Nov-1999
C:Accession: A48030; A60164; A35483; C35483; A60432; A47507; S34329
R:Language: F.; Morales, M.; de la Salle, C.; Cazenave, J.P.; Clemenson, K.J.; Shimomura, T.
J. Biol. Chem. 268, 20801-20807, 1993
A:Title: Cloning and characterization of the gene encoding the human platelet glycoprotein V.
A:Reference number: A48030; MUID:94012616; PMID:8407908
A:Accession: A48030
A:Molecule type: DNA

Residues: 1-560 <LA2>

A:Cross-references: EMBL:Z23091; NID:9312501; PID:CAAB0637.1; PID:9312502
R:Shimomura, T.; Fujimura, K.; Maehama, S.; Takemoto, M.; Oda, K.; Fujimoto, T.; Oyama, A.
Blood 75, 2349-2356, 1990
A:Title: Rapid purification and characterization of human platelet glycoprotein V: the a

A:Accession: A60164; MUID:90275263; PMID:2350580

A:Molecule type: Protein

A:Residues: 365-384, 'X', 386-390, 'X', 392-395, 'X', 397, 188-208, 'T', 210, 27-50, 'X', 52-53, 174-
'XX', 108, 'T', 61-72, 'TK', 75-77, 'V', 56-57, 'G', 479-487, 'X', 489-498, 'X', 500, 'X', 502-503, 'X',
R:Roth, G.J.; Church, T.A.; McMullen, B.A.; Williams, S.A.
Biochem. Biophys. Res. Commun. 170, 153-161, 1990

A:Title: Human platelet glycoprotein V: a surface leucine-rich glycoprotein related to a
A:Reference number: A35483; MUID:90321220; PMID:2372284
A:Accession: A35483

A:Molecule type: Protein

A:Residues: 145-166, 'T', 168-169, 'X', 171-172 <RO7>

A:Note: this proteolytic fragment was designated peptide M392

A:Accession: B35483

A:Molecule type: Protein

A:Residues: 121-129, 'W', 131-135, 466-468, 'X', 470 <RO2>

A:Note: this material was designated peptide M393 but may contain two peptides

A:Accession: C35483

A:Molecule type: Protein

A:Residues: 253-266, 'H', 268-272, 'X', 274-279, 'T', 281-284, 'T', 286 <RO3>

A:Note: this proteolytic fragment was designated peptide M401

R:Zafar, R.S.; Walz, D.A.
Thromb. Res. 53, 31-44, 1989

A:Title: Platelet membrane glycoprotein V: characterization of the thrombin-sensitive gl
A:Reference number: A60432; MUID:89162331; PMID:2922700
A:Accession: A60432

A:Molecule type: Protein

A:Residues: 477-478, 'FX', 481-485, 'E', 487, 'V', 489-492, 'NQ', 495, 'E', 497-498 <ZAF>
R:Hikey, M.J.; Hagen, F.S.; Yagi, M.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 8327-8331, 1993

A:Title: Human platelet glycoprotein V: characterization of the polypeptide and the rela

A:Reference number: A47507; MUID:93391348; PMID:7690959
A:Accession: A47507

A>Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-560 <RES>

A:Cross-references: GB:L11238; NID:9388759; PID:AAA03069.1; PID:9388760

C:Comment: This platelet membrane protein is a substrate for thrombin.

C:Comment: The amino end of the intact protein is blocked.

C:Comment: This protein is absent in Bernard-Soulier syndrome.

C:Genetics:

A:Gene: GDB:GP5

A:Cross-references: GDB:230236; OMIT:173511

A:Map position: 5pter-5qter

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

C:Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane prot

Query Match 14.5%; Score 323; DB 2; Length 560;

Best Local Similarity 31.5%; Pred. No. 2,5e-16;
Matches 113; Conservative 38; Mismatches 162; Indels 46; Gaps 12;

QY 49 NNESSVP---ISLPSTORFLQNNL-----ITTLAPRTG 81
Db 180 NNLTTHLPKGLGAQAKLERLLHSNRVLSLDSGLNSGALTEILOFHHNRHSIAPGAFD 239
QY 82 --SNLLTLWLFNSNLTSTYPTGFRHLQALEEDLDGNRHRLSLPDTFOGLERLOSLHLY 139
Db 240 RLPLNLSLTLSNNHLAFLPSALFLSHNLTLLTFEN-PLALPGLVLGKMGLOELMKN 298
QY 140 RCQLSSPENTFRGLVSLQYLYLOENSLH-LQDDLADLANLSHFLHGNRLLTTEHV 198
Db 299 RTQLRTLPAARNLRSRLRYLTGVTLSPLSALPQGAFOGIGELQVALHSNGLTALPGGL 358
QY 199 FPGLSLDRLLHGNRLQGVHRAAFRGLSRLLTTLFNNLSLALSGEALADPSLEFLR 258
Db 359 LGLGLKRLQVSLRRNRRLRLPRALFRNLSSLESYVDHQLSTLPGLDYGALRLEVL 418
QY 259 NANPWACDCRARPLWAFQR-ARYSSD-VTCAPPPEKGRDLRALREAPDQACPPAP 315
Db 419 GHSNWRCCDGLPFLGLWRHLGLVGEPPRCAGPAGHAGLPWALPGDAE-CP--GP 475
QY 316 TRPGSRARKNSSNNHLYGAEEAGAPPADPSTLYRDLPAEDSRGQGDAPTEDDYWG 374
Db 476 RGPFRPADSSSE--APVHPALAPNSSEPMWAO-PVTYTKGOD-----HSPFWGFY 525

RESULT 10

T42218
slit-1 protein homolog - rat

N:Alternate names: MEGF4 protein

C:Species: Rattus norvegicus (Norway rat)

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2002

R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998

A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motif

A:Reference number: Z14126; MUID:98360089; PMID:9693030

A:Accession: T42218

A>Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1531 <NNA>

A:Cross-references: EMBL:AB011530; NID:93449289; PIDN:BA32460.1; PID:93449290

A:Experimental source: strain Sprague-Dawley; brain

C:Genetics:

A:Gene: MEGF4

C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein

Query Match 13.9%; Score 310; DB 2; Length 1531;
Best Local Similarity 27.3%; Pred. No. 7.9e-15;
Matches 93; Conservative 51; Mismatches 138; Indels 59; Gaps 7;

QY 17 LILMLALPLAPSPCMCTCYSSPTVSCQANNFSSVPLSLPSQORFLQNNLIRTLR 76
Db 20 LILMAAWRLGATACALCTGT-TVDCHGTLQALIKNTIPRNTRELELN----- 69
QY 77 PGTFGSNLLTLWLFNSNLTSTYPTGFRHLQALEEDLDGNRHRLSLPDTFOGLERLOSL 136
Db 70 -----GNNTIRIHKNDPAGLKKQLRVLODMNQ-IGAVERGAFFDMKLERL 114
QY 137 HLRYCOLSSLPGNIFRGVLSQYLYLOENSLHLLQDDLADLANLSHFLHGNRLLTTE 196
Db 115 RLNRNOLQVLELPLQNNALSRDLSENS-----LDAYPR 150
QY 197 HYFRGLGSIDRLHGNRLQGVHRAAFRGLSRLLTTLFNNLSLALSGEALADPSLEFL 256
Db 151 KAFRGYTDLKNQLQDNQISCIIEGAFRALRGVEVLTLLNNNTTTPVSSFNMPKLRFF 210
QY 257 RLNANWACDCRARPLWAFQARVSSDVTCATPPEKGRDLRALREAPDQACPPAP 309
Db 211 RLHSNHLFDCHLAWLSQMLRQRLPTIGLFTQCSGASLGLNVAEYQKSEFSGGGEAA 270

RESULT 2

A41915

Insulin-like growth factor-binding complex acid-labile chain precursor - human
 N:Alternate names: Acid-labile Subunit (ALS)
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999

C:Accession: A41915
 R:Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.
 Mol. Endocrinol. 6, 870-876, 1992

A:Title: Structure and functional expression of the acid-labile subunit of the insulin-1
 A:Reference number: A41915; MUID:9235025; PMID:1379671

A:Accession: A41915
 A:Status: preliminary
 A:Molecule type: mRNA; protein

A:Residues: 1-605 <LEO>
 A:Cross-references: GB:M66826; NID:g184807; PIDN:AAA36047.1; PID:g184808

A:Experimental source: liver
 A:Note: sequence extracted from NCBI backbone (NCBIRP:110171)

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
 F:1-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:112/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:113-116/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:117-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:111-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F:291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
 F:315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F:339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
 F:363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F:411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
 F:435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F:459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
 F:483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>

F:507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>

Query Match 15.7%; Score 349; DB 2; Length 605;
 Best Local Similarity 28.8%; Pred. No. 3.3e-18;

Matches 109; Conservative 33; Mismatches 115; Indels 122; Gaps 8;

```

QY 47 QANNFSSVPLSLPSTQRLFLQNNLIRTLRPTFGS----- 82
DB 234 KANVFVOL-----PRLQKLYLDRLNLAIAVAPGAFGLKALRWLDSHNVAQLLEDTPFG 288
DB 83 --NLTLTLMFSSNLSTIYGTFRHLQALELDGDNRLRSLEPDTFGSLRQLSLHYR 140
DB 289 LGLRLVRLSHNATASLRLPTEFKDLFEELQDGNHR-TROLAERSFEGQLQLEVLTLDH 347
QY 141 QQLSSLPNGIFRGVLSLOYLYQENSLSLHODDLFADLANLSLFLHGNRLLEHVR 200
DB 348 NQLQEVKAGAFGLTNVAVNNISGNCNRLRPROVGRGLKLSLHESGSLRIRPHFT 407
QY 201 GIGSLDRLLHGNRLQGV-----HRAERGLSRLTILYLF 235
DB 408 GLSGLRLRLKQNGVLGEEQSIMGLLELDTLSNQLTHLPHR-LRQGLCKLEYLLS 466
QY 236 NNSLASLPEALADL-----P 251
DB 467 RNRLELPAADLGLQRAFWLDVSHNRLEALPNLSLALGLRLRYSLRNNSLRFTTPOP 526
QY 252 SLFELRLANPACDCRRAPFLAM-----FORARVSSSD-----VTCATP 291
DB 527 GLERLMLTEGNPWDCGPKALRDFALQNPASAVRFVQALICGDDCQPPATYNNITCASP 586
QY 292 PERQGDRLALREADFOAC 310
DB 587 PEVYGLDLRDLSEAHFAPC 605

```

RESULT 3

A53531

oncofetal trophoblast glycoprotein 5T4 precursor - human
 N:Alternate names: oncofetal antigen 5T4
 C:Species: Homo sapiens (man)
 C:Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 05-Nov-1999

C:Accession: A53531
 R:Myers, K.A.; Rahl-Sund, V.; Davison, M.D.; Young, J.A.; Cheater, A.J.; Stern, P.L.
 J. Biol. Chem. 269, 9319-9324, 1994

A:Title: Isolation of a cDNA encoding 5T4 oncofetal trophoblast glycoprotein. An anti
 A:Reference number: A53531; MUID:94179356; PMID:8132670

A:Accession: A53531
 A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-420 <MEY>
 A:Cross-references: EMBL:D29083; NID:g435654; PIDN:CAA82324.1; PID:g435655

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
 C:Keywords: duplication; glycoprotein; transmembrane protein

F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-420/Product: oncofetal trophoblast glycoprotein 5T4 #status predicted <MAT>

Query Match 14.9%; Score 331.5; DB 2; Length 420;
 Best Local Similarity 31.9%; Pred. No. 4.2e-17;

Matches 102; Conservative 26; Mismatches 141; Indels 51; Gaps 10;

```

QY 23 ALPLAAPSCPMILCTGSSPPYSCANNFSSVPLSLPSTQRLFLQNNLIRTLRPTFG- 81
DB 54 AQPRLPDGCPALCESEARATVKCNRLTEVPDLPAYVNLFLTGQLAVLPAGAFAR 113
QY 82 ----SNLTLTLMFSSNLSTIYGTFRHLQALELDGDNRLRSLEPDTFG----- 129
DB 114 RPLLELALNLSGSRIDDEVRAQAFENLPSLRQDLDSHN-PLADLSPFARSGSNASVAP 172
QY 130 --LERQLSLHY-----RCQLSSLPNGIF-----RGVLSLOYLYQENSLSLHODDLFAD 177
DB 173 SPLVELLNHTVPEDEQNSFEQGVYVALLAGRALQGLRLLELSNHLFLPRDLAQ 232
QY 178 LANLSLFLHGNRLRLTEHVRFGLSLDRLLHGNRLQGVHRAERGLSRLTILYLFNN 237
DB 233 LPSLRHLDLSSNLSVLSLYVSFRNLTHLESILHEDNALKYLH-----NG 276
QY 238 SLASLPGEALADLPSLEFLRLANPACDCRRAPFLAMPRARV-SSSVYTCATPERO 295
DB 277 TLAELOG-----LRHIVF-LDNNPWCDCMHADMTWLKETEYVQCKRLTCAYBEKMR 330
QY 296 GRDRLALREADFOACPPAP 315
DB 331 NRVLLEINSADLD-CDPILP 349

```

RESULT 4

JC6128

Insulin-like growth factor binding complex acid labile chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 23-Mar-1997 #sequence_revision 09-May-1997 #text_change 05-Nov-1999

C:Accession: JC6128
 R:Boisclair, Y.R.; Seto, D.; Hsieh, S.; Hurst, K.R.; Ooi, G.T.
 Proc. Natl. Acad. Sci. U.S.A. 93, 10028-10033, 1996

A:Title: Organization and chromosomal localization of the gene encoding the mouse ac1
 A:Reference number: JC6128; MUID:96413591; PMID:8816745

A:Accession: JC6128
 A:Molecule type: DNA

A:Residues: 1-603 <ROI>
 A:Cross-references: GB:U66900; NID:g1621612; PIDN:AAB17270.1; PID:g1621613

C:Comment: This protein is a serum protein and it is of the ternary complex in the ph
 A:Gene: als

A:Map position: 17

Query Match 14.8%; Score 329.5; DB 2; Length 603;
 Best Local Similarity 35.3%; Pred. No. 9.1e-17;

Matches 95; Conservative 34; Mismatches 99; Indels 41; Gaps 5;

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 12:07:28 ; Search time 46 Seconds
(Without alignments)
877.749 Million cell updates/sec

Title: US-09-972-546-2
Accession: 2326
Sequence: 1 MLGCLRLQAPASACLLM.....LSAGLPPLCLLLVPHHL 420

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

1 number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	15.9	605	2 JC5239	insulin-like growth
2	349	15.7	605	2 A41915	insulin-like growth
3	331.5	14.9	420	2 A53531	oncocyte trophobol
4	329.5	14.8	603	2 JC6128	insulin-like growth
5	326.5	14.7	603	2 JC1282	insulin-like growth
6	324.5	14.6	1469	2 B36655	slit protein 2 pre
7	324.5	14.6	1480	2 A36655	chondroderin pre
8	323	14.5	361	2 A33860	platelet membrane
9	323	14.5	560	2 A60164	slit-1 protein hom
10	310	13.9	1531	2 T42218	MEGF5 protein - ra
11	293	13.2	1523	2 T13953	neuronal leucine-r
12	286.5	12.9	707	2 JC7763	leucine-rich alpha
13	280.5	12.6	312	1 NBH0A2	glial cell membran
14	278.5	12.5	1091	2 A58532	lysine carboxypept
15	273.5	12.3	536	2 A34901	decorin precursor
16	265	11.9	360	2 S06280	peroxidase - fru1
17	261.5	11.7	1535	2 S42799	garp precursor - h
18	258.5	11.6	662	2 S42799	decorin precursor
19	253	11.4	359	1 NBH0C8	orphan G protein-c
20	253	11.4	907	2 JE0176	decorin - rabbit
21	249	11.2	360	2 JG0193	G protein-coupled
22	248.5	11.2	907	2 JG0193	hypothetical prote
23	243.5	10.9	789	2 T28714	hypothetical prote
24	243.5	10.9	1355	2 T28715	secreted leucine-r
25	241.5	10.8	1025	2 T42626	proline-arginine-
26	240.5	10.8	382	2 T39068	decorin precursor
27	239.5	10.8	357	2 S24317	decorin precursor
28	236.5	10.6	354	2 S29145	cell-surface molec
29	232.5	10.4	682	2 A49121	

30	232.5	10.4	682	2 A43318	connectin precursor
31	231.5	10.4	1112	2 T10504	disease resistance
32	231.5	10.4	1389	2 T13852	gene wheeler prote
33	229	10.3	626	1 NBH01A	platelet glycoprot
34	228.5	10.3	1385	2 T13887	slit protein - fru1
35	226.5	10.2	354	2 A55454	decorin precursor
36	226.5	10.2	562	2 T34319	hypothetical prote
37	220	10.1	594	2 T23841	hypothetical prote
38	220	9.9	594	2 T23841	hypothetical prote
39	217.5	9.8	610	2 T19938	disease resistance
40	211	9.5	1016	2 T30553	hypothetical prote
41	206.5	9.3	224	2 T32185	hypothetical prote
42	206.5	9.3	603	2 T24315	hypothetical prote
43	205.5	9.2	1066	2 T15864	hypothetical prote
44	204.5	9.2	961	2 T23395	hypothetical prote
45	204.5	9.2	1134	1 A29944	chaoptin precursor

ALIGNMENTS

RESULT 1
JC5239
insulin-like growth factor acid-labile chain - baboon
C:Species: Papio sp. (baboon)
C:Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
C:Accession: JC5239
R:Delnathy, P.; Baxter, R. C.
Biochem. Biophys. Res. Commun. 227, 897-902, 1996
A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-
A:Reference number: JC5239; MUID:97040714; PMID:8886027
A:Contents: liver
A:Accession: JC5239
A:Molecule type: mRNA
A:Residues: 1-605
C:Comment: This factor is structurally related to proinsulin and have insulin-like m

Query Match 15.9%; Score 353; DB 2; Length 605;
Best Local Similarity 28.4%; Pred. No. 1.7e-16;
Matches 122; Conservative 40; Mismatches 130; Indels 138; Gaps 9;

QY	2	LPGLRLQA-----PASACLLMLALPLAPSCPMCTCYSPPTVSCQANFSS	53
DB	193	LGCLRLVLAGRLAYIQALPFGSLAELELDLSNRLAI-----KANVFAQ	240
QY	54	VPUSPPSTQRLFNNTLRTPGFS-----NLTTL	87
DB	241	L-----PRLQKLYLDRLIAAVALGFLKALRWLDLSHNRVAGLDETPGGLGLRVL	295
QY	88	WLFSNNLSTTYGTRRHQALELDGDN-----H-----	118
DB	296	RSHNIAISLRPTFEDLFLLELDGHNRIQLAERSTEGIGLEVLTLDHNOGEYKV	355
QY	119	-----LRSLEPTFGGLRIQSLIHYRCOLSSLPGNITFGVLSIOTL	160
DB	356	GATLGITNAVMNLSCNCRNLPEQVFRGLKLSHIEGSLGIRPTFAGLSGLRL	415
QY	161	LYOENSLHLODDLPDLANLSHLPLHGNRLRLHEHVRGSGSDRLRLHGNRLOGVHR	220
DB	416	FLKDNGLVGIIEOSLWGLLELLELDLSNQLTHLPQLEGLKLEYLLSHNRLELPA	475
QY	221	AARGLISRLTIYLFNNLSLSPGALADJ-----PSLEPLRLNA	260
DB	476	DALGPIQRAFWDVSHNRLEALPGLSLASIGRLRYLNRNNSLRPTTPOPLERLMEIG	535
QY	261	NPWACDGRARPLWAM-----FORARVSSD-----VTCAPEPRGRGRDR	300
DB	536	NPMDCSPLKALRDFALQNSAVPREVOALCEGDDQPPVYTYNNITCASPREVAGLDR	595
QY	301	ALREAPDQAC 310	
DB	596	DLGEAHFAPC 605	

FT	REPEAT	172	195	LRR 5.
FT	REPEAT	197	219	LRR 6.
FT	REPEAT	407	430	LRR 7.
FT	REPEAT	431	454	LRR 8.
FT	REPEAT	455	478	LRR 9.
FT	REPEAT	480	502	LRR 10.
FT	REPEAT	503	526	LRR 11.
FT	REPEAT	528	550	LRR 12.
FT	DOMAIN	712	723	POLY-GLY.
SO	SEQUENCE	977 AA;	109005 MW;	3C936B7E0003DF54 CRC64;

Query Match 13.1% Score 291; DB 1; Length 977;
Best Local Similarity 29.7% Pred. No. 2,3e-13;
Matches 98; Conservative 37; Mismatches 133; Indels 62; Gaps 13;

OY	31	CPMLCTC-----YSSPPYVSCQANNSSVPLSLPPSTQRLFLQNNLIRTLRPGTGSNLLT	86
Db	373	CPTGCTCNLHINDGLTFVNCKERGFNNISELP-----RP-----LNKK	412
OY	87	LMLFSNNLSTIYPTGFRHLALBELDGDNRHLRSLPDTFGGLERLQSLHLRCQLSSL	146
Db	413	LYLSSNLIQKTYRSDFWNFSSLDLHLGNNR-ISTYODGAFINLPNLKSLFLNGNDIEKL	471
OY	147	PGNIFRGLVSLQYLYLOENSLHLQDDCFADLANLSHLFGNRLRLTEHVFRLGSLD	206
Db	472	TPGMFRGLQSLHLXLYFEFNVIREIQPAFSLMPLKLLFLNNLIRTLPTDAFAGT-SLA	530
OY	207	RLLLHGNRLOGVHRAAFRGLSRLLTILYLFNNSLASLPGEALADLPSEFLRLNANPMACD	266
Db	531	RLNLRKN-----YFLYL-----PVAG-VLEHLNAIVQIDLNENPMDCI	567
OY	267	CRAAPLWAMFORARVSS----SDVTCATPPEROGDRLALREADFOACPPAAPTREGSRA	322
Db	568	CDLVPFKQWIE--TISVSVVGVDYLCRSPENLTHRDVRI-ELEV-LCPBMLHAVAPAGES	623
OY	323	RGNSSNHLGVAAAG-----APPADPSTL	347
Db	624	PAQPDHSLIGAPTSAPYEFSPPGSPVPL	653

Search completed: April 4, 2003, 12:08:34
Job time : 18 secs


```

OY 283 SSDVTCAPPERQGHDLRAL--READFOCAPPAAP--TRPGSR--ARGNSSNHLTYGVAEA 337
DB 392 POSTLCAEPDQLRLPVREVPREMTDCLPLISPRSPPLQVAGSGSMVLEHRAALAE- 450
OY 338 GAPPADPSFLYDLDAEASRSGOGDAPTEDDYMGYGEGEDQRCQMP---GAACQAP 394
DB 451 -----PEPEIYW-----VTPAGLRPLTPAHAGRCRYRP 478

OY 395 DS-----RGPALACL 405
DB 479 EGTLELRVTAEAGL 494

RESULT 13
CHAD_HUMAN STANDARD: PRT: 359 AA.
ID CHAD_HUMAN STANDARD: PRT: 359 AA.
AC O15335; Q96R35;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
15-JUN-2002 (Rel. 41, Last sequence update)
Chondroadherin precursor (Cartilage leucine-rich protein).
CHAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE=98008928; PubMed=9344663;
RT Grover J., Chen X.-N., Korenberg J.R., Roughley P.J.;
RT "The structure and chromosome location of the human chondroadherin
RT gene (CHAD).";
RL Genomics 45:379-385(1997).
RN 12
RP SEQUENCE FROM N.A.
RA MEDLINE=21413956; PubMed=11445564;
RT Maansson B., Wengelin C., Moergelin M., Saxne T., Helinggaard D.;
RT "Association of chondroadherin with collagen type II.";
RL J. Biol. Chem. 276:32883-32888(2001).
CC -1- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and
CC osteoblasts. This binding is mediated (at least for chondrocytes
CC and fibroblasts) by the Integrin alpha(2)beta(1). May play an
CC important role in the regulation of chondrocyte growth and
CC proliferation (By similarity).
CC -1- SUBUNIT: Mostly monomeric (By similarity). Interacts with collagen
CC type II.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
CC -1- TISSUE SPECIFICITY: Present in chondrocytes at all ages.
CC -1- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
CC (SLRPs) FAMILY. CLASS IV SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: U96769; AAC13410.1;
CC EMBL: U96767; AAC13410.1; JOINED.
CC EMBL: U96768; AAC13410.1; JOINED.
CC EMBL: AF371328; AAK51556.1;
CC Genew; HGNC:1909; CHAD.
CC MIM: 602178;
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000483; LRR_Cterm.
CC InterPro: IPR000372; LRR_Nterm.
CC InterPro: IPR003592; LRR_out.
CC InterPro: IPR003591; LRR_Lyp.
CC Pfam; PF00560; LRR; 10.

```

```

DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00370; LRR; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYRP; 4.
KW Signal; Leucine-rich repeat; Repeat; Glycoprotein;
KW Extracellular matrix.
FT SIGNAL 1 22
FT CHAIN 23 359
FT REPEAT 50 73
FT REPEAT 74 97
FT REPEAT 99 121
FT REPEAT 122 145
FT REPEAT 147 169
FT REPEAT 171 193
FT REPEAT 194 217
FT REPEAT 218 241
FT REPEAT 243 266
FT REPEAT 267 290
FT REPEAT 292 317
FT DISULFID 23 38
FT DISULFID 304 346
FT DISULFID 306 326
FT CARBOHYD 144 144
FT CONFLICT 114 114
FT CONFLICT 166 166
SQ SEQUENCE 359 AA; 40487 MW; 9A318D0A15C157A0 CRC64;

Query Match 14.1%; Score 313; DB 1; Length 359;
Best Local Similarity 28.6%; Pred. No. 1; Re-15;
Matches 104; Conservative 47; Mismatches 141; Indels 72; Gaps 11;

OY 19 LMLALPL-----AASCPMLCTCYSPPTVSCQANNFSSVPLSLPSTQRLFLQNNLI 72
DB 5 MLLSLUGLALGLPALACGPONCHSDLOHVICDKYGLQKIP-KVSEKTKLNLQRRNF 63
OY 73 RFLRPGTFS--NLTLWLFENNLSITYPGFRLHQLLEFLDGDNNHLSLEPDTQGL 130
DB 64 PVLAANSFRAPNLVSLHLOQCIREVAAGAFRGKOLITLYLSHN-DIVVRAQAFDDL 122
OY 131 ERLDSHLHYRQLSSLPNGT-----FRQLVSLQVLYLOENS 166
DB 123 TELVLYLDHNKVKVELPRGLSLPLVNLPIQLNNKIRELRAGPQCAKDLRWLYSENA 182
OY 167 LHLDDLEFDLANLS-----HLFLHGRRLRLTEINFRGL 202
DB 183 LSLDLPALDQVENLAFHYDRNQLSSPSAALSKLRVVEELKLSHPKSTIPDNAFQSF 242
OY 203 GS-LDRLLHGNRLQGVHRAAFRGLSLRTLYLENNSLASLPGEALADLPSTLEFLRLNAN 261
DB 243 GRVLETMLDNTNLEKFSDCAGFLGVTTLKHVHLENNRLNQLPSNFPFD-SLETTLATLNN 300
OY 262 PMACDCRAPPMLAMFORARVSSSDVTCAPPEROG---RDLRLRADQACPPAATRP 318
DB 301 PMWCTQQLRLRLRWLE-AKASRPDATCASPANKFGOHIRDTDAFRSCKP-----PTKR 352
OY 319 GSRA 322
DB 353 SKRA 356

RESULT 14
PLIB_AGRBL
ID PLIB_AGRBL STANDARD: PRT: 331 AA.
AC 093233;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phospholipase A2 inhibitor subunit B precursor (PLI-B).
OS Agkistrodon blomhoffii sinicus (Chinese mamsuhi) (Gloydinus
OS blomhoffii sinicus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


CC -1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U96626; AAC39963.1; -
DR EMBL: BC012672; AAH12672.1; -
DR MGD: MGI:1096866; Chnu.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003591; LRR_Lyp.
DR Pfam: PF00560; LRR_10.
DR Pfam: PF01463; LRRCT; 1.
DR Pfam: PF01462; LRRNT; 1.
DR SMART: SM00082; LRRNT; 1.
DR SMART: SM00013; LRRCT; 1.
DR SMART: SM00369; LRR_Lyp; 5.
DR Signal: Leucine-rich repeat; Repeat; Glycoprotein;
KW Extracellular matrix.
FT SIGNAL 1 20
FT CHAIN 21 358
FT REPEAT 49 72
FT REPEAT 73 96
FT REPEAT 97 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 217 240
FT REPEAT 242 265
FT REPEAT 266 290
FT REPEAT 294 317
FT DISULFID 22 37
FT DISULFID 303 345
FT DISULFID 305 325
FT CARBOHYD 143 143
SQ SEQUENCE 358 AA; 40348 MW; 6A062FCEBF84A078 CRC64;

Query Match 14.4%; Score 320.5; DB 1; Length 358;
Best local similarity 29.2%; Pred. No. 5 3e-16;

Matches 105; Conservative 50; Mismatches 138; Indels 67; Gaps 12;

DB 17 LLMALLPLAASCPMLCTCTSSPTVSCQANFSSVPLSLPSTORFLQNNLIRFLR 76
9 LVLALILPL-ALAACPQNHCHGDLOHVICDKVGLQKIP-KVSETTKLLNQRNFPVLA 66
OY 77 PCFFGS--NLTLMLPSSNLSITVPTFRHLQALELDLGDNRHLRSLPDTFGGLERLQ 134
DB 67 ANSFRTMPLVSHLOHCHIREVAAAGAFGLKOLLYLYLSHN-DIRVLRAGAFDILTET 125
OY 135 SMLHRCQLSSLPNGT-----FRGLVSLQVLYLOENSLHL 170
DB 126 YLYLHNKVSSELPRLGLSPVLNFIQLNNKIRELRAGAGADLWVLSENAISSL 185
OY 171 QDDLEADLANLS-----HLFNGNRLRLTEHVEFGGLS-L 205
DB 186 QPGSIDVDENLAKFLDKNQSSYPSAALSILRVVEELKLSHPNPKSTPDVAFOFSGRYL 245
OY 206 DRLLLHGNRLQGVHRAFRGLSRLTILYLFNNSLASLDEGALADLPSEFLRLANPAPAC 265
DB 246 ETLWMDNTNLEKFSDAEFGVTTLKHVHLNDNRLNQLPSSFPFD--NLETTLTNNPDKC 303
OY 266 DCRAPLMAMPARAVSSDVTGATPPRPG--RDALRLREADFOACPPAPTPRPSRA 322
DB 304 TCQLRGLRLWLE-AKASRPDNTCSSPAKFGQRIKIDTDLR-----SC--KSPTRRSKKA 355

RESULT 11

CHAD_RAT	STANDARD:	PRT:	358 AA.
ID CHAD_RAT			
AD 070210;			
DT 15-JUN-2002 (Rel. 41, Created)			
DT 15-JUN-2002 (Rel. 41, Last sequence update)			
DT 15-JUN-2002 (Rel. 41, Last annotation update)			
DE Chondroadherin precursor (Cartilage leucine-rich protein).			
GN CHAD.			
OS Rattus norvegicus (Rat).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX NCBI_TaxID=10116;			
RN 11)			
RP SEQUENCE FROM N.A.			
RC TISSUE=Chondrosarcoma;			
RA MEDLINE=98129774; PubMed=9461555;			
RT Shen Z., Gancheva S., Maansson B., Helnegard D., Sommarin Y.,			
RL Biochem. J. 330:549-557(1998).			
CC -1- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and			
CC osteoblasts. This binding is mediated (at least for chondrocytes			
CC and fibroblasts) by the Integrin alpha(2)beta(1). May play an			
CC important role in the regulation of chondrocyte growth and			
CC proliferation (By similarity).			
CC -1- SUBUNIT: Mostly monomeric (By similarity).			
CC -1- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).			
CC -1- TISSUE SPECIFICITY: Present in femoral head and rib cartilage, as			
CC well as in tendon. Detected in bone marrow.			
CC -1- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEGLYCANS			
CC (SLRPs) FAMILY. CLASS IV SUBFAMILY.			
CC -1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL: AF004953; AAC40060.1; -			
DR InterPro: IPR001611; LRR.			
DR InterPro: IPR000483; LRR_Cterm.			
DR InterPro: IPR000372; LRR_Nterm.			
DR InterPro: IPR003591; LRR_Lyp.			
DR Pfam: PF00560; LRR_10.			
DR Pfam: PF01463; LRRCT; 1.			
DR Pfam: PF01462; LRRNT; 1.			
DR SMART: SM00082; LRRCT; 1.			
DR SMART: SM00013; LRRNT; 1.			
DR SMART: SM00369; LRR_Lyp; 5.			
KW Signal: Leucine-rich repeat; Repeat; Glycoprotein;			
KW Extracellular matrix			
FT SIGNAL 1 20			
FT CHAIN 21 358			
FT REPEAT 49 72			
FT REPEAT 73 96			
FT REPEAT 97 120			
FT REPEAT 121 144			
FT REPEAT 146 168			
FT REPEAT 170 192			
FT REPEAT 193 216			
FT REPEAT 217 240			
FT REPEAT 242 265			
FT REPEAT 266 289			
FT REPEAT 291 316			
FT DISULFID 22 37			
FT DISULFID 303 345			
FT DISULFID 305 325			
FT CARBOHYD 143 143			
SQ SEQUENCE 358 AA; 40403 MW; 630946F96F384857 CRC64;			

CC TISSUE=platelet; PubMed=2372284;
 CC MEDLINE=90321220; McMullen B.A., Williams S.A.;
 CC Roth G.J., Church T.A., McMillen B.A., Williams S.A.;
 CC "Human platelet glycoprotein V: a surface leucine-rich glycoprotein
 CC related to adhesion.";
 CC Blochem. Biophys. Res. Commun. 170:153-161(1990).
 CC -1- FUNCTION: THE GPIIb-V-III COMPLEX FUNCTIONS AS THE VON WILLEBRAND
 CC FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT
 CC PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO
 CC - INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A
 CC CRITICAL INITIATING EVENT IN HEMOSTASIS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: PLATELETS AND MEGAKARYOCYTES.
 CC -1- PIV: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: L11238; AAA03069.1; -;
 CC DR EMBL: Z23091; CAA80637.1; -;
 CC DR HSSP: P16473; LXUM.
 CC DR Genew: HGNC:4443; GP5.
 CC DR MIM: 173511;
 CC DR InterPro: IPR001611; LRR.
 CC DR InterPro: IPR000483; LRR_Cterm.
 CC DR InterPro: IPR000372; LRR_Nterm.
 CC DR InterPro: IPR003592; LRR_out.
 CC DR InterPro: IPR003591; LRR_typ.
 CC DR Pfam: PR00560; LRR: 13.
 CC DR Pfam: PF01463; LRCT: 1.
 CC DR PRINTS: PRO0019; LEORICHRPT.
 CC DR SMART: SM00370; LRR: 1.
 CC DR SMART: SM00082; LRCT: 1.
 CC DR SMART: SM00013; LRRNT: 1.
 CC DR SMART: SM00369; LRR_TYP: 9.
 CC KW Platelet; Transmembrane; LRR; Typ: 9.
 CC Repeat; Leucine-rich repeat; Cell adhesion; Signal.
 CC FT CHAIN 1 16
 CC FT GN 1
 CC FT DOMAIN 17 560
 CC FT TRANSMEM 524 544
 CC FT DOMAIN 545 560
 CC FT REPEAT 73 96
 CC FT REPEAT 97 120
 CC FT REPEAT 122 144
 CC FT REPEAT 145 168
 CC FT REPEAT 169 192
 CC FT REPEAT 194 216
 CC FT REPEAT 217 240
 CC FT REPEAT 241 264
 CC FT REPEAT 266 288
 CC FT REPEAT 289 312
 CC FT REPEAT 314 337
 CC FT REPEAT 338 361
 CC FT REPEAT 362 385
 CC FT REPEAT 386 409
 CC FT REPEAT 409 499
 CC FT CARBOHYD 51
 CC FT CARBOHYD 181
 CC FT CARBOHYD 243
 CC FT CARBOHYD 267
 CC FT CARBOHYD 298
 CC FT CARBOHYD 312
 CC FT CARBOHYD 385
 CC FT CARBOHYD 499
 CC FT CARBOHYD 73
 CC FT CONFLICT 109
 CC FT CONFLICT 130

FT CONFLICT 136 138 GID -> PG (IN REF. 3).
 FT CONFLICT 209 209 L -> I (IN REF. 2).
 FT CONFLICT 267 267 N -> H (IN REF. 3).
 FT CONFLICT 327 327 N -> I (IN REF. 2).
 FT CONFLICT 478 478 P -> G (IN REF. 2).
 FT CONFLICT 509 509 P -> D (IN REF. 2).
 FT CONFLICT 509 509 P -> D (IN REF. 2).
 SQ SEQUENCE 560 AA: 60959 MW: BICDB04F8AF7115 CRC64;
 Query Match 14.5%; Score 323; DB 1; Length 560;
 Best Local Similarity 31.5%; Pred. No. 6e-16; Indels 46; Gaps 12;
 Matches 113; Conservative 38; Mismatches 162;
 QY 49 NNESSVP---LSLPSRQRLFLNNL-----IRTLRPGTEG 81
 DB 180 NMLTHPKGLGAGAKERLLLSNRLVSLDSGLNSLGLATFLQFRHNRHSIARGAD 239
 QY 82 --SNLTLWLFSSNLSITYPGTFRHLQALEELDGDNRHLRSLEPDTFGLELSHLX 139
 DB 240 RLRLSLSLTISRNLAVLPALFLSHNLTLLTFEN-PLAELPGVLFEGMGLQELWLN 298
 QY 140 RQGLSLPGNIFRGVLSQVLYIQENSLH-LQDDLEFADANTSHFLHGRLRLTTEHY 198
 DB 299 RTQRLTPAAAFRNLISRLRYGLVLSRLSALPGAGFQIGELQVLAHNGTLALPDGL 358
 QY 199 FRGLSLDRLLHGNRLQGVHRAFRGLSRITLILFNNSLASIPGEALADPSLEFLRL 258
 DB 359 LRGLKLRQVSLRRNRLRALPRALFRNLSSLESVOLHNOLETLPGDFVGLRLEEVLL 418
 QY 259 NANPWACDCRRAPLMANFOR--AVSSSD-VTCATPERGGRDRLAREADFOCAPPA 315
 DB 419 GNSWRKDCGGLPGLGWLRQHLGLVGGEPPRCAGPGAHAGLPALPGDAE-CP-CP 475
 QY 316 TRPGSRANGSSNNHLYVAAGAPPAAPDPTLYRDLPAEDSRGSGDAFTEDDYMGY 374
 DB 476 RQPPRPAAADSSE--AVVHPALAPNSEPWWAQ-PVTTGKGD-----HSPFWGFY 525
 RESULT 10
 CHAD_MOUSE
 ID CHAD_MOUSE STANDARD; PRT; 358 AA.
 AC 055226;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chondroadherin precursor (Cartilage leucine-rich protein).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OC NCBI_TaxID=10090;
 OX 11
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98126439; PubMed=9465299;
 RA Landgren C., Beier D.R., Faessler R., Heinegaard D., Sommarin Y.;
 RT "The mouse chondroadherin gene: characterization and chromosomal
 RT localization.";
 RL Genomics 47:84-91(1998).
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RA Strausberg R.;
 RU Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and
 CC osteoblasts. This binding is mediated (at least for chondrocytes
 CC and fibroblasts) by the integrin alpha(2)beta(1). May play an
 CC important role in the regulation of chondrocyte growth and
 CC proliferation (By similarity).
 CC -1- SUBUNIT: Mostly monomeric. Interacts with collagen type II (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
 CC -1- TISSUE SPECIFICITY: Cartilage.
 CC -1- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN
 CC (SLRPS) FAMILY. CLASS IV SUBFAMILY.

```

RN [2]
RP SEQUENCE OF 25-55 AND 77-97.
RC TISSUE-BONE;
RA MEDLINE-95113864; PubMed-7814406;
RX Hu B., Coulson L., Moyer B., Price P.A.;
RT "Isolation and molecular cloning of a novel bone phosphoprotein
RT related in sequence to the cystatin family of thiol protease
RT inhibitors."
RL J. Biol. Chem. 270:431-436(1995).
CC -1- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and
CC osteoblasts. This binding is mediated (at least for chondrocytes
CC and fibroblasts) by the integrin alpha2beta(1). May play an
CC important role in the regulation of chondrocyte growth and
CC proliferation.
CC -1- SUBUNIT: Mostly monomeric. Interacts with collagen type II (by
CC similarity).
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- TISSUE SPECIFICITY: Cartilage.
CC -1- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
CC (SLRPs) FAMILY. CLASS IV SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U08018: AAA21330.1; -.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR_10.
DR Pfam: PF01462; LRRNT; 1.
DR Pfam: PF01463; LRRCT; 1.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 5.
DR Signal: Leucine-rich repeat; Repeat: Glycoprotein;
KW Extracellular matrix.
FT CHAIN 1 24 OR 23 (IN SOME ISOFORM(S)).
FT CHAIN 25 361 CHONDRODHERIN.
FT REPEAT 52 75 CHONDRODHERIN, MINOR FORM.
FT REPEAT 76 99 LRR 1.
FT REPEAT 100 123 LRR 2.
FT REPEAT 124 147 LRR 3.
FT REPEAT 149 171 LRR 4.
FT REPEAT 173 195 LRR 5.
FT REPEAT 196 219 LRR 6.
FT REPEAT 220 243 LRR 7.
FT REPEAT 245 268 LRR 8.
FT REPEAT 269 292 LRR 9.
FT REPEAT 294 319 LRR 10.
FT DISULFID 25 40 LRR 11.
FT DISULFID 306 348 BY SIMILARITY.
FT DISULFID 308 328
FT CARBOHYD 146 146
FT CONFLICT 25 25 O-LINKED (POTENTIAL).
FT CONFLICT 29 29 C -> Y (IN REF. 2).
FT CONFLICT 31 31 C -> H (IN REF. 2).
FT CONFLICT 40 40 C -> L (IN REF. 2).
FT CONFLICT 40 40 S -> R (IN REF. 2).
FT CONFLICT 52 52
SQ SEQUENCE 361 AA; 40884 MM; DA79DC98AD3DD1D8 CRC64;
Query Match
Best Local Similarity 14.6%; Score 324; DB 1; Length 361;
Matches 105; Conservative 51; Mismatches 140; Indels 68; Gaps 12;
QY 17 LILMLLALPL-----AAPSCPMILCTCYSSPTVSCQANNFSSVPLSLPSTQRLFLQNN 70

```

```

DB 5 MLLXLSLGLASLLPALAACPNCHCHSDLOHVIDKVGLOKIP-KYSEKRLNLORN 63
QY 71 LIRLRGPGFGS--NLTTLMPFSNNSTIYPGFRRQLALEBLDGNHRLSLPDPFQ 128
DB 64 NPFVLATNSFRAMPNLVSHLOHCOIREVAAGAFNGLKOLIVLYLSHN-DIVLBAAGFD 122
QY 129 GLERQSLHLRYCQSLSLPGNI-----PRLVSLQYLYQIE 164
DB 123 DLTELYLYLDHKNVLEPRGLSLPLVNFPIQLNNKRIELRSAGFQGAKDLRWLYLSE 182
QY 165 NSLLHLODDLFADLANLSHLFLHGRN-----LRLTEHYFR 200
DB 183 NSSLSLQPGALDDVENTLAKFYLDNRQLSSPSAALSKLRYVEELKLSHNPILKSIDPNAFQ 242
QY 201 GLGS-LDRLLHGNRLQGVHRAAFRLSLRTLYLEFNNSLASLPGELADLPSEFLALN 259
DB 243 SFGRYLETMLDNTNLEKSDSAFLGVTTLKHYHLENNRHLQLPSPFPD-SLETLTLYT 300
QY 260 AMPWACDCARFLAMFORARVSSDYTCAPPERQGDRLALREAD-FQACPPAAPTRP 318
DB 301 NNPWACTQQLRGLRRLWE-AKTSRPDATCASPAPKFRGQH---IRDTAFRCG--KPTKR 354
QY 319 GSRA 322
DB 355 SKKA 358
RESULT 9
GPV_HUMAN
ID GPV_HUMAN STANDARD; PRT; 560 AA.
AC P40197;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Platelet glycoprotein V precursor (GPV) (CD42D).
OS GPs.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-Lung;
RA MEDLINE-93391348; PubMed-7690959;
RA Hickey M.J., Hagen F.S., Yagi M., Roth G.J.;
RT "Human platelet glycoprotein V: characterization of the polypeptide
RT and the related Ib-V-IX receptor system of adhesive, leucine-rich
RT glycoproteins."
RL Proc. Natl. Acad. Sci. U.S.A. 90:8327-8331(1993).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE-Platelet;
RA MEDLINE-94012616; PubMed-8407908;
RA Lanza F., Morales M., de la Salle C., Cazenave J.-P., Clemetson K.J.,
RA Shimomura T., Phillips D.R.;
RT "Cloning and characterization of the gene encoding the human platelet
RT glycoprotein V. A member of the leucine-rich glycoprotein family
RT cleaved during thrombin-induced platelet activation."
RL J. Biol. Chem. 268:20801-20807(1993).
RN 13
RP PARTIAL SEQUENCE.
RC TISSUE-Platelet;
RA MEDLINE-90275263; PubMed-2350580;
RA Shimomura T., Fujimura K., Maehama S., Takemoto M., Oda K.,
RA Fujimoto T., Oyama R., Suzuki M., Ichihara-Tanaka K., Titani K.,
RA Kuramoto A.;
RT "Rapid purification and characterization of human platelet
RT glycoprotein V: the amino acid sequence contains leucine-rich
RT repetitive modules as in glycoprotein Ib."
RL Blood 75:2349-2356(1990).
RN 14
RP PARTIAL SEQUENCE.

```

DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01187; EGF_CA; 2.
 DR PROSITE; PS00025; LAM_G_DOMAIN; 1.
 DR Neurogenesis; Glycoprotein; Signal: Alternative splicing;
 KW EGF-like domain; Repeat; Leucine-rich repeat.
 FT SIGNAL
 FT CHAIN 1 1480
 FT REPEAT 37 1480 SLIT PROTEIN.
 FT REPEAT 99 122 LRR 1.
 FT REPEAT 123 146 LRR 2.
 FT REPEAT 148 170 LRR 3.
 FT REPEAT 171 194 LRR 4.
 FT REPEAT 195 218 LRR 5.
 FT REPEAT 220 246 LRR 6.
 FT REPEAT 321 344 LRR 7.
 FT REPEAT 345 368 LRR 8.
 FT REPEAT 369 392 LRR 9.
 FT REPEAT 394 417 LRR 10.
 FT REPEAT 417 440 LRR 11.
 FT REPEAT 522 545 LRR 12.
 FT REPEAT 546 569 LRR 13.
 FT REPEAT 570 593 LRR 14.
 FT REPEAT 595 617 LRR 15.
 FT REPEAT 618 641 LRR 16.
 FT REPEAT 643 666 LRR 17.
 FT REPEAT 678 701 LRR 18.
 FT REPEAT 720 743 LRR 19.
 FT REPEAT 745 764 LRR 20.
 FT REPEAT 765 788 LRR 21.
 FT REPEAT 790 812 LRR 22.
 FT REPEAT 813 836 LRR 23.
 FT REPEAT 838 861 LRR 24.
 FT REPEAT 907 944 EGF-LIKE 1.
 FT DOMAIN 946 983 EGF-LIKE 2.
 FT DOMAIN 985 1022 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1024 1062 EGF-LIKE 4.
 FT DOMAIN 1064 1100 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1111 1149 EGF-LIKE 6.
 FT DOMAIN 1152 1325 LAMININ G-LIKE.
 FT DOMAIN 1353 1392 EGF-LIKE 7.
 FT DOMAIN 1409 1480 CRCK
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 998 998 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1159 1159 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1175 1175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1292 1292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 911 922 BY SIMILARITY.
 FT DISULFID 916 932 BY SIMILARITY.
 FT DISULFID 934 943 BY SIMILARITY.
 FT DISULFID 950 961 BY SIMILARITY.
 FT DISULFID 955 971 BY SIMILARITY.
 FT DISULFID 973 982 BY SIMILARITY.
 FT DISULFID 989 1001 BY SIMILARITY.
 FT DISULFID 995 1010 BY SIMILARITY.
 FT DISULFID 1012 1021 BY SIMILARITY.
 FT DISULFID 1028 1041 BY SIMILARITY.
 FT DISULFID 1035 1050 BY SIMILARITY.
 FT DISULFID 1052 1061 BY SIMILARITY.
 FT DISULFID 1068 1079 BY SIMILARITY.
 FT DISULFID 1073 1088 BY SIMILARITY.
 FT DISULFID 1090 1099 BY SIMILARITY.
 FT DISULFID 1115 1125 BY SIMILARITY.

FT DISULFID 1120 1137 BY SIMILARITY.
 FT DISULFID 1139 1146 BY SIMILARITY.
 FT DISULFID 1157 1168 BY SIMILARITY.
 FT DISULFID 1357 1368 BY SIMILARITY.
 FT DISULFID 1362 1380 BY SIMILARITY.
 FT DISULFID 1382 1391 BY SIMILARITY.
 FT DISULFID 1409 1443 BY SIMILARITY.
 FT DISULFID 1423 1457 BY SIMILARITY.
 FT DISULFID 1434 1473 BY SIMILARITY.
 FT DISULFID 1438 1475 BY SIMILARITY.
 FT DISULFID 1442 1479 BY SIMILARITY.
 FT VARSPLIC 1394 1404 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 1480 AA; 165752 MW; F9D5925FC170B1C3 CRC64;
 Query Match 14.6%; Score 324.5; DB 1; Length 1480;
 Best Local Similarity 23.5%; Pred. No. 1.5e-15;
 Matches 101; Conservative 53; Mismatches 122; Indels 153; Gaps 6;
 OY 31 CPMCTCYSSPPVSCQANNSSVPLSPSTORPLRPLONNLITLRPTFG---SNLLTLW 88
 DB 73 CPRVCSC--TGLNVDCSHRGLTSVPRKISADVERLELOGNNLTVIETDFQRLTKRLMLQ 130
 OY 89 LFSNNLSTIYPTGRHRLQALEELDLGDN-----RHLSLEPD----- 125
 DB 131 LTDMQHTIERNSPQDLVSLERLDSNNVITTVGRVFKQAQSLRSIQDNNQITCDEH 190
 OY 126 TFOGRLRSLHLRCLSLPLGNIFRGLVSLQYLOENSL----- 167
 DB 191 AFKGLVEILITLNNNNITSLPHNIFGLRLRLSDNPFACDCHLSWLSRFLRSATR 250
 OY 168 -----LHLD----- 172
 DB 251 LAPYRCSPSQLGQNVADLHDFKCSGLTEHAPMECAENSCPHPCACADSYDRE 310
 OY 173 -----DLFADLA 179
 DB 311 KSLSVAPTLDDDTVDVLEQNFTELPRKSSFRRLRIDLSNNNISRLAHDAISGLK 370
 OY 180 NLSHLFNGNRLRLTEHVPRGLSLDRLLHGNRLQCVHNAARGLSRLITLTFNNSL 239
 DB 371 QLTTLVLGNKIKIDLPSSGVFKGLSLRLILNANEISQIRDAFRDLSLSLYDNI 430
 OY 240 ASLPEALADLPSEFLNLANPACDCRAPLWAMFORAVSSDYCATPERQGRDL 299
 DB 431 QSLANGTPEAMSKMTVHLAKNPFICDCNKLWLAADYLHKNPLETSGACESPKNRHRRI 490
 OY 300 RALREADPQ 308
 DB 491 ESLREKFK 499
 RESULT 8
 CHAD_BOVIN STANDARD; PRT; 361 AA.
 ID CHAD_BOVIN
 AC 027972;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Chondroadherin precursor (cartilage leucine-rich protein) (38 kDa bone
 DE protein).
 DE CHAD.
 OS Bos taurus (Bovine).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 OX [1]
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP TISSUE=Cartilage;
 RC MEDLINE=94342341; PubMed=8063792;
 RA Neame P.J., Sommarin Y., Boynton R.E., Heinegaard D.;
 RT "The structure of a 38-kDa leucine-rich protein (chondroadherin)
 RT isolated from bovine cartilage";
 RL J. Biol. Chem. 269:21547-21554(1994).


```

CC EMBL: S46785; AAB23770.2; -
DR PIR: JCI282; JCI282.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_Out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00560; LRR_18.
DR Pfam: PF01463; LRRNT; 1.
DR Pfam: PF01463; LRRCT; 1.
DR Pfam: PR00019; LEURICHPT.
DR SMART: SM00370; LRR; 5.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 9.
KM Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT CHAIN 1 23
FT SIGNAL 24 603
INSULIN-LIKE GROWTH FACTOR BINDING
PROTEIN COMPLEX ACID LABILE CHAIN.
REPEAT 52 73
REPEAT 74 96 LRR 1.
REPEAT 98 120 LRR 2.
REPEAT 121 144 LRR 3.
REPEAT 146 168 LRR 4.
REPEAT 169 192 LRR 5.
REPEAT 194 216 LRR 6.
REPEAT 217 240 LRR 7.
REPEAT 242 264 LRR 8.
REPEAT 266 288 LRR 9.
REPEAT 289 312 LRR 10.
REPEAT 313 336 LRR 11.
REPEAT 337 360 LRR 12.
REPEAT 361 384 LRR 13.
REPEAT 386 408 LRR 14.
REPEAT 409 432 LRR 15.
REPEAT 433 456 LRR 16.
REPEAT 458 478 LRR 17.
REPEAT 479 504 LRR 18.
REPEAT 506 529 LRR 19.
REPEAT 543 566 LRR 20.
REPEAT 566 586 LRR 21.
CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 603 AA; 66811 MW; DCD7637D94A5037C CRC64.

Query Match 14.7% Score 326.5; -DB 1; Length 603;
Best Local Similarity 35.1% Pred. No. 3.7e-16;
Matches 94; Conservative 35; Mismatches 99; Indels 41; Gaps 5;

```

```

RESULT 7
SLIT_DROME STANDARD: PRT: 1480 AA.
ID SLIT_DROME
AC P24014;
DT 01-MAR-1992 (rel. 21, Created)
DT 01-MAR-1992 (rel. 21, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE SLIT protein precursor.
GN SLIT.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91099655; PubMed=2176636;
RA Rothberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
RT "Slit: an extracellular protein necessary for development of midline
RT glia and commissural axon pathways contains both EGF and LRR
RT domains."
RL Genes Dev. 4:2169-2187(1990).
CC - FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND
CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR
CC MATRIX MOLECULES.
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.
CC - SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
CC - SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
CC - SIMILARITY: CONTAINS 1 C-TERMINAL CISTINE KNOT-LIKE (CKCK) DOMAIN.
CC - SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X53959; CA37910.1; -
CC PIR: A36655; A36655.
CC HSSP: P00740; 1EDM.
CC Flybase: FBgn0003425; slit.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000359; Cys_knot.
DR InterPro: IPR000561; EGF_Like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR000359; LRR_Out.
DR InterPro: IPR003591; LRR_Typ.
DR InterPro: IPR001791; Laminin_G.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00008; EGF_7.
DR Pfam: PF00054; Laminin_G; 1.
DR Pfam: PF00560; LRR_16.
DR Pfam: PF01462; LRRNT; 4.
DR Pfam: PF01463; LRRCT; 4.
DR SMART: SM00041; CT; 1.
DR SMART: SM00179; EGF_Ca; 2.
DR SMART: SM00001; EGF_Like; 5.
DR SMART: SM00370; LRR; 4.
DR SMART: SM00082; LRRCT; 4.
DR SMART: SM00013; LRRNT; 4.
DR SMART: SM00369; LRR_Typ; 9.
DR SMART: SM00282; Lamg; 1.

```


OY 102 SHLEHGRRLTLTEHVEFGSLDRLLLHGNRLQGVHRAAFRGISRLTILFNNSLAS 241
 DB 342 RVLAVHTNALBELPDLALGRLROVSLRHNRLRALPRTLFRNLSSLVYQLEHNO LKT 401
 OY 242 LPGEALADLPSEFLRLNANPWACDCRARPPLAMFQRRARVSSDYTCATPEPREGDRLA 301
 DB 402 LRGDVFALPOLTRVLGHNPWLDCD---GLMPFLQMLRHNLLELGRDEPQCNGPSSRA 458
 OY 302 -----LREADFOAC-----PPAAPT-----PPGSR-----ARGNSSN 329
 DB 459 SLTEWELLOGD-QMCPSSRGRLPPDPPTENALKADPPTQRRPSSQSMANVOLVANGESP DN 517
 OY 330 HLY 332
 DB 518 RPY 520

RESULT 4
 GPV_MOUSE STANDARD: PRT: 567 AA.
 008742:

15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Platelet glycoprotein V precursor (GPV) (CD42D).
 GN GP5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RA MEDLINE=97275136; PubMed=9129030;
 RA Ravanat C., Morales M., Azorsa D.O., Moog S., Schuhler S.,
 RA Grunert P., Loew D., van Dorsselaer A., Cazenave J.-P., Lanza F.:
 RT Gene cloning of rat and mouse platelet glycoprotein V:
 RT Identification of megakaryocyte-specific promoters and demonstration
 RT of functional thrombin cleavage.".
 RL Blood 89:3253-3263(1997).
 CC -1- FUNCTION: THE GPIIb-IIIa COMPLEX FUNCTIONS AS THE VON WILLEBRAND
 CC FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT
 CC PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO
 CC INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A
 CC CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: 269595; CA93441.1; .
 DR HSSP: P09661; IAGN.
 DR MGD: MGI:1096363; GP5.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_LYP.
 DR Pfam: PF00560; LRR_12.
 DR Pfam: PF01463; LRRCT_1.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR_1.
 DR SMART: SM00082; LRRCT_1.
 DR SMART: SM00013; LRRNT_1.
 DR SMART: SM00369; LRR_TYP_10.
 DR Platelet; Transmembrane; Glycoprotein; Blood coagulation;
 DR Repeat; Leucine-rich repeat; Cell adhesion; Signal.
 KW

FT SIGNAL 1 16
 FT CHAIN 17 567
 FT DOMAIN 17 522
 FT TRANSMEM 523 543
 FT DOMAIN 544 567
 FT REPEAT 73 96
 FT REPEAT 97 120
 FT REPEAT 122 144
 FT REPEAT 145 168
 FT REPEAT 170 192
 FT REPEAT 194 216
 FT REPEAT 217 240
 FT REPEAT 241 264
 FT REPEAT 266 288
 FT REPEAT 289 312
 FT REPEAT 314 337
 FT REPEAT 361 385
 FT REPEAT 362 385
 FT REPEAT 386 409
 FT CARBOHYD 51 51
 FT CARBOHYD 67 67
 FT CARBOHYD 181 181
 FT CARBOHYD 243 243
 FT CARBOHYD 298 298
 FT CARBOHYD 312 312
 FT CARBOHYD 385 385
 SQ SEQUENCE 567 AA: 63467 MW: C48643AAV3967AVD CRC64;
 Query Match 15.4%; Score 343.5; DB 1; Length 567;
 Best Local Similarity 30.7%; Pred. No. 2.1e-17;
 Matches 119; Conservative 41; Mismatches 148; Indels 79; Gaps 10;

OY 49 NFESSVPLSPSTQRLPL-----QNNLRLTRPPTFGS--NLTLMLFSNNLSTYPTQTF 102
 DB 156 NOLSLPANLSSRLBELKLDLSRNNLN-LPKGLGAQVRLKELLYSNOLTSVDSGLL 214
 OY 103 RHLQALEELDGDNRHLRLSLRPTFOGLERLQSLHLYGQSLSPGNFFRGLYSLOYLYL 162
 DB 215 SNLGLTELRL-ERNHLSVAPGAFDRGLNSLSLTSGLNLSLELPPALFLHYSSVSRLL 273
 OY 163 QENSLLHODDLFPDLATLSHLFLHG-----
 DB 274 FENPLELPDVLFGMGLRELMLNGHLSTLPPAAFRNLSGLOTGLTRNPRSLALPRG 333
 OY 189 -----NRLRLTEHVEFGSLDRLLLHGNRLQGVHRAAFRGISRLTILY 233
 DB 334 VFOGLRELVLGLHTNMLAELRDALRGILGRVSLRHNRLRALPRTLFRNLSSLESVO 393
 OY 234 LFNNSLASLPREALADLPSEFLRLNANPWACDCRARPPLAMFQRRARVSSDYTCATPE 293
 DB 394 LEHMOLETLPGDVFALPOLQYVLGINPWLDCD---GLMRFQMLRHNHDPDILGRDEPQ 450
 OY 294 RQGRDLRA-----LREADFOACPPAAPTRPGSRARCNSSNHLGYGAEGAPADPSTL 347
 DB 451 CRGPEPRASLSFWELLOGD-PMCPD-----PRSLPLDPPTENALKADPPTQRRPSSQ 504
 OY 348 YRDLPADSRKRGCGADPTEDDYGWGY 374
 DB 505 WAQLVAR-----GESPNRLYWGILY 524

RESULT 5
 ALS_MOUSE STANDARD: PRT: 603 AA.
 ID ALS_MOUSE
 AC P70389;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein complex acid labile chain
 DE precursor (ALS).
 GN IGFBP3 OR ALS OR ALBS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

FT REPEAT 265 288 LRR 10.
FT REPEAT 289 312 LRR 11.
FT REPEAT 313 336 LRR 12.
FT REPEAT 337 360 LRR 13.
FT REPEAT 361 384 LRR 14.
FT REPEAT 386 408 LRR 15.
FT REPEAT 409 432 LRR 16.
FT REPEAT 433 456 LRR 17.
FT REPEAT 458 480 LRR 18.
FT REPEAT 482 504 LRR 19.
FT REPEAT 505 530 LRR 20.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 605 AA; 66034 MW; F5562A23CBE918F6 CRC64;

```

```

Query Match 15.7%; Score 349; DB 1; Length 605;
St. Local Similarity 28.8%; Pred. No. 9,2e-18;
Matches 109; Conservative 33; Mismatches 115; Indels 122; Gaps 8;

```

```

OY 47 QANNESSVLSLPSTQRLFLQNNLIRLPPTFGS----- 82
234 KANVEVOL-----PRLOKLYLDNRLLIAVAAPGAGLKLRLMLDSHNFAGLLEDTPFG 288
83 --NLITLWFSNNLSTIYGFGRHQALEBDLGDNRHLRSLEPTFGQLERLQSLHYR 140
289 LGLRLRLSHNIAIARLRTFKDLHLELDLGNHR -IRQLAERSFEGIGOLEVLTLDH 347
OY 141 COLSSIPGNTFRGLVSLQYLYLOENSLHLODDLPADLANSHLPLHGRKLLTFHYR 200
348 NLOEVKAGAGLGLTFVAAVMNLGNCRLMLPEQVFRGLKLSLHSGSLGRIRPHIT 407
OY 201 GLGSLDRLLHGNRLQV-----HRAAFGLSLRTLLYLF 235
408 GLSGRLRLFKDNGLVGIEQSLMGLAELELDLTSNQLTHLPHR -LFGIGLKEVLLLS 466
OY 236 NNSIASLPGBALADL-----P 251
467 RNRLAELPADALGPLQAFWLDVSHNRLEALPNSLAPLGRLYLSLRNNSLRTFPPOP 526
OY 252 SLEPLRLNAPWACDCARLYNAM-----FORARYSSD-----VTCATP 291
527 GLERLWLEGPMPCGCPKALRDPALONPSAVRFVQALCEGSDCCPATTYNNITCASP 586
292 PERGRDLRALREADFQAC 310
587 PEVVGIDLRLDLSAHPAC 605

```

RESULT 3
 GPV_RAT STANDARD; PRT; 567 AA.
 AC 008770;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Platelet glycoprotein V precursor (GPV) (CD42D).
 GN GP5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OC NCBI_TaxID=10116;
 OX NCBI [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=97275136; Pubmed=9129030;
 RA Ravant C., Morales M., Azorsa D.O., Moog S., Schuhler S.,
 Grunert P., Loew D., van Doorselaer A., Gazeave J.-P., Lanza F.;
 "Gene cloning of rat and mouse platelet glycoprotein V:
 identification of megakaryocyte-specific promoters and demonstration

```

RT of functional thrombin cleavage."
RL Blood 89:3253-3262(1997).
CC -1- FUNCTION: THE GPIIb-V-IX COMPLEX FUNCTIONS AS THE VON WILLEBRAND
CC FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT
CC PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO
CC INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A
CC CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z69594; CAA93440.1; -.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000483; LRR_Cterm.
CC InterPro: IPR000372; LRR_Nterm.
CC InterPro: IPR003592; LRR_Out.
CC InterPro: IPR003591; LRR_Typ.
CC Pfam: PF00560; LRR. 13.
CC Pfam: PF01463; LRRCT. 1.
CC PRINTS: PR00019; LEURICHRPT.
CC SMART: SM00370; LRR. 2.
CC SMART: SM00082; LRRCT. 1.
CC SMART: SM00013; LRRNT. 1.
CC SMART: SM00369; LRR_Typ. 10.
CC Platelet; Transmembrane; Glycoprotein; Blood coagulation;
CC Repeat; Leucine-rich repeat; Cell adhesion; Signal.
CC SIGNAL 1 16
CC CHAIN 17 567
CC DOMAIN 17 522
CC TRANSMEM 523 543
CC DOMAIN 544 567
CC REPEAT 73 96
CC REPEAT 97 120
CC REPEAT 122 144
CC REPEAT 145 168
CC REPEAT 169 192
CC REPEAT 194 216
CC REPEAT 217 240
CC REPEAT 241 264
CC REPEAT 266 288
CC REPEAT 289 312
CC REPEAT 314 337
CC REPEAT 338 361
CC REPEAT 362 385
CC REPEAT 387 409
CC REPEAT 409 498
CC CARBOHYD 51 51
CC CARBOHYD 181 181
CC CARBOHYD 243 243
CC CARBOHYD 298 298
CC CARBOHYD 312 312
CC CARBOHYD 365 365
CC CARBOHYD 498 498
SQ SEQUENCE 567 AA; 63344 MW; CA10708ED03707F CRC64;

```

```

Query Match 15.6%; Score 347; DB 1; Length 567;
St. Local Similarity 36.0%; Pred. No. 1,2e-17;
Matches 109; Conservative 27; Mismatches 127; Indels 40; Gaps 9;

```

OY 65 LFLQNNLIRLPPTFGS--NLITLWFSNNLSTIYGFGRHQALEBDLGDNRHLRSL 122
 223 LRLERNHLRSIADGADSLNLSLTLGSLSLPDLFHVSWLTRLLEN -PLEEL 281
 OY 123 EPTFGQLERLQSLHYRQLSSIPGNTFRGLVSLQYLYLOENSLH -LODDLPADLANL 181
 282 PEVLFGMAQLRELWNGHLRLPLPAAPRNLSGLQTLRNPDLISALPFGMFLHTEL 341

```

FT REPEAT 53 73 LRR 1.
FT REPEAT 74 96 LRR 2.
FT REPEAT 98 120 LRR 3.
FT REPEAT 121 144 LRR 4.
FT REPEAT 146 168 LRR 5.
FT REPEAT 169 192 LRR 6.
FT REPEAT 194 216 LRR 7.
FT REPEAT 217 240 LRR 8.
FT REPEAT 242 264 LRR 9.
FT REPEAT 265 288 LRR 10.
FT REPEAT 289 312 LRR 11.
FT REPEAT 313 336 LRR 12.
FT REPEAT 338 360 LRR 13.
FT REPEAT 361 384 LRR 14.
FT REPEAT 386 408 LRR 15.
FT REPEAT 409 432 LRR 16.
FT REPEAT 433 456 LRR 17.
FT REPEAT 458 480 LRR 18.
FT REPEAT 482 504 LRR 19.
FT REPEAT 505 530 LRR 20.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 605 AA: 66110 MM: 9D71994625F23652 CRC64:

```

Query Match 15.9%; Score 353; DB 1; Length 605;
 Best Local Similarity 28.4%; Pred. No. 4.7e-18;

Matches 122; Conservative 40; Mismatches 130; Indels 138; Gaps 9;

```

OY 2 LFGRRLLA-----PASALLMLLALPLAASCPMLCTCYSSPTVSCQANFSS 53
DB 193 LGLRLVLAAGNLAVALQPALESGLAELRELDLSNALRAI-----KANVFAQ 240
OY 54 VPLSPSTQRLFLQNNLIRTLRPTFGS-----NLTL 87
DB 241 L-----PRLOKLYLDRNLAAVAAPGAFGLKMLRWLDLSHNVAAGLEPTFGGLGLRYL 295
OY 88 WLFSSNLSTIPTGFRHLQALEPLDGDNR-----H----- 118
DB 296 RLSHNAIASLRPTFEDLHFLELDLQGNHRIQALERSFEGIGOLEVLTLDHNOQDEVK 355
OY 119 -----LRSLEPTFGQLERLOSLLHYRCQLSSIPGIFRGLSYLYL 160
DB 356 GAFGLTNVAVMNLSCNCRLNPEQVFRGLGLSHLHLESGSLGRTPRTFAGSLGRYL 415
OY 161 YQENSLHLQDDLPADLANLSHLHGNRLRLTEHVFRLGSLDRLLHGNRLQCYHR 220
DB 416 FLKDNGLVIGEOSLWGLLELDLTSNQLNHLPHQLFOGLKLEYLLSHNRLAELPA 475
OY 221 AAFRGLSRITLYLFNNSLASLPGELADL-----PSLEFLRLNA 260
DB 476 DALGFLQRAFWLDVSHNRLAELPGSLSLSLGLRLRLNNSLRTPPOPLELMLLEG 535
OY 261 NPMACDCRRAPLMAW-----FORARVSSD-----VTCATPEPGRGLR 300
DB 536 NPMDCSPLKALRDPAALNPSAVPRFVAICGDDCQPPVYTYNNTCASPEVAGDLR 595
OY 301 ALREADFOAC 310
DB 596 DLGEAHFAPC 605

```

RESULT 2
 ID ALS_HUMAN STANDARD: PRT: 605 AA.

AC P35858:
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein complex acid labile chain

```

DE precursor (ALS).
GN IGFALS OR ALS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RA MEDLINE=92357025; PubMed=1379671;
RX Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
RT "Structure and functional expression of the acid-labile subunit of
RL the insulin-like growth factor-binding protein complex.";
RM Mol. Endocrinol. 6:870-876(1992).
RP SEQUENCE FROM N.A.
RA Frankland J.;
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
RN (3)
RP SEQUENCE OF 28-35.
RX MEDLINE=89308584; PubMed=2473065;
RA Baxter R.C., Martin J.L., Beniac V.A.;
RT "High molecular weight insulin-like growth factor binding protein
RT complex. Purification and properties of the acid-labile subunit from
RL human serum.";
RL J. Biol. Chem. 264:11843-11848(1989).
CC -1- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
CC IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
CC IGF-1 OR IGF-II AND IGFBP-3.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: M86826; AAA36047.1;
DB EMBL: AL031724; CAC36078.1;
DB PIR: A41915; A41915.
DB Genew: HGNC:5468; IGFALS.
DB MIM: 601489;
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR_19.
DR Pfam: PF01462; LRRNT; 1.
DR Pfam: PF01463; LRRCT; 1.
DR PRINTS: PR00019; LEURCHRP.
DR SMART: SM00370; LRR_2.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_TYP; 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; signal.
FT SIGNAL 1 27
FT CHAIN 28 605
FT REPEAT 53 73 INSULIN-LIKE GROWTH FACTOR BINDING
FT REPEAT 74 96 PROTEIN COMPLEX ACID LABILE CHAIN.
FT REPEAT 98 120 LRR 1.
FT REPEAT 121 144 LRR 2.
FT REPEAT 145 168 LRR 3.
FT REPEAT 170 192 LRR 4.
FT REPEAT 193 216 LRR 5.
FT REPEAT 217 240 LRR 6.
FT REPEAT 242 264 LRR 7.
FT REPEAT 264 LRR 8.
FT REPEAT 264 LRR 9.

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 12:07:27 ; Search time 14 Seconds
(without alignments)
1244.290 Million cell updates/sec

Title: US-09-972-546-2

Perfect score: 2226
Sequence: 1 MLTGLRLAPASACLLM.....LSAGLPPLICLLVPHNL 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	15.9	605	1	ALS_PAPHA
2	349	15.7	605	1	ALS_HUMAN
3	347	15.6	567	1	GPV_RAT
4	343.5	15.4	567	1	GPV_MOUSE
5	329.5	14.8	603	1	ALS_MOUSE
6	326.5	14.7	603	1	ALS_RAT
7	324.5	14.6	1480	1	SLIT2_DROME
8	324	14.6	361	1	CHAD_BOVIN
9	323	14.5	560	1	GPV_HUMAN
10	320.5	14.4	358	1	CHAD_MOUSE
11	318.5	14.3	358	1	CHAD_RAT
12	314	14.1	713	1	GAC1_HUMAN
13	313	14.1	359	1	CHAD_HUMAN
14	309.5	13.9	331	1	PLTB_ACKBL
15	291	13.1	977	1	Y848_HUMAN
16	287	12.9	646	1	FLR1_HUMAN
17	283	12.7	951	1	LGR4_RAT
18	282	12.7	660	1	LGR4_HUMAN
19	282	12.6	951	1	LGR4_HUMAN
20	280.5	12.6	312	1	A2GL_HUMAN
21	275.5	12.4	649	1	FLR3_HUMAN
22	274.5	12.3	966	1	Y918_HUMAN
23	273.5	12.3	536	1	CBP8_HUMAN
24	265	11.9	360	1	PGS2_BOVIN
25	261	11.7	360	1	PGS2_SHEEP
26	259	11.6	360	1	PGS2_CANFA
27	258.5	11.6	662	1	GARP_HUMAN
28	258	11.6	360	1	PGS2_PIG
29	256	11.5	360	1	PGS2_HORSE
30	253	11.4	907	1	PGS2_HUMAN
31	253	11.4	907	1	LGR5_HUMAN
32	249	11.2	360	1	PGS2_RABIT
33	248.5	11.2	907	1	LGR5_MOUSE

34	240.5	10.8	356	1	PGS2_COTUA
35	240.5	10.8	382	1	PRIP_HUMAN
36	239.5	10.7	357	1	PGS2_CHICK
37	239	10.7	378	1	PRIP_MOUSE
38	237	10.6	757	1	LGR7_HUMAN
39	236.5	10.6	354	1	PGS2_RAT
40	234.5	10.5	381	1	PRIP_BOVIN
41	232.5	10.4	377	1	PRIP_RAT
42	232.5	10.4	682	1	CONN_DROME
43	229	10.3	626	1	GPBA_HUMAN
44	226.5	10.2	354	1	PGS2_MOUSE
45	218	9.8	754	1	LGR8_HUMAN

ALIGNMENTS

RESULT 1
ALS_PAPHA ID ALS_PAPHA STANDARD; PRT; 605 AA.
AC 002833;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
GN IGFALS OR ALS.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Papio.
CX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97040714; PubMed=8886027;
RA Delhanty P., Baxter R.C.;
RT "The cloning and expression of the baboon acid-labile subunit of the insulin-like growth factor binding protein complex."
RT Biochem. Biophys. Res. Commun. 227:897-902(1996).
CC -!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL: S83462; -; NOT ANNOTATED_CDS.
DR InterPro: IPR001611; LRR
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF000560; LRR: 19
DR Pfam: PF01462; LRRNT: 1.
DR Pfam: PF01463; LRRCT: 1.
DR PRINTS: PR00019; LEURICHREP.
DR SMART: SM00370; LRR: 2.
DR SMART: SM00082; LRRCT: 1.
DR SMART: SM00013; LRRNT: 1.
DR SMART: SM00369; LRR_TYP: 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 605
FT INSULIN-LIKE GROWTH FACTOR BINDING
FT PROTEIN COMPLEX ACID LABILE CHAIN.

Fri Apr 4 12:29:25 2003

us-09-972-546-2.rai

OY	114	GDNRLHSLEEDPTQGGIERLOSGLHYFCGOLSSLEGNTFRGLVSLQYLXLOENSLHLHDD	173
Dd	106	-QGSMRLSLEEQALLGIDNLYLTLHELRNRKLNLAVGFTPTTPSTASTISLSNNLLGRLEE	164
OY	174	LFAADLANISHFLFCHGNRIRLTTEHVPRFGSLSDRLDLHGHRLOCVHBAAFRGTSRLTIY	233
Dd	165	LFOGLSHMDLNLGMSNSTLVLPDPIVPGCLGHLVELLAGNKLTLYLQPALFCGLGEELRD	224
OY	234	LFNNSLASLPEALADLPSELEFLRLNAN	261
Dd	225	LSRNALRSVKANVFVHLPRLOKLTLDRN	252

```

US-08-487-072A-50
: RESULT 10
: Sequence 50 Application US/08487072A
: Patent No. 6423884
: GENERAL INFORMATION:
: APPLICANT: Mochly-Rosen, Daria
: APPLICANT: Ron, Dorit
: TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
: TITLE OF INVENTION: Thereof
: NUMBER OF SEQUENCES: 265
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morlison & Foerster
: STREET: 2000 Pennsylvania Avenue, NW
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20006-1812
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487,072A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: MURASHIGE, KATE H.
: REGISTRATION NUMBER: 29,959
: REFERENCE/DOCKET NUMBER: 2550-0025.20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 887-1500
: TELEFAX: (202) 887-0763
: INFORMATION FOR SEQ ID NO: 50:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 603 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: INDIVIDUAL ISOLATE: Insulin-like growth factor bind
: INDIVIDUAL ISOLATE: pro. complex-1a1, Fig. 33
US-08-487-072A-50

```

```

Query Match      14.7% Score 326.5 DB 4: Length 603:
Best Local Similarity 35.1% Pred No 1.9e-22:
Matches 94; Conservative 35; Mismatches 96; Indels 41; Gaps 5

OY   12 PASACLLMLLAL-----PLAA---PSCPMLCTC---YSSPTVSCQANFSS 53
      | | : | | | | | | | | | | | | | | | | | | | | | | |
Db    8 PALVVLLAFWVALGCHLOGTDPGASADAEGQCVCVACTSHDDYTDDELVSFCSSKLNTH 67

OY   54 VPLSPSTORLFILNNLTIRLPSTFGSLLTLTFMSNLSTIYPGPFRHQALEIDL 113
      | : | | | | | | | | | | | | | | | | | | | | | | |
Db    68 LPDIPVSTR-----ALWDGGNLSITSMAFQNLSLDPLFL 105

OY   114 GDNRRLSLRPPTFGLERLOSLLHYRCQLSSLPCGNIRGVLSLOYLYQENSLHLLODD 173
       ||||| | | | | | | | | | | | | | | | | | | | |

```

Db 106 -qSMWLRSLPEQALLGIGONLYLHLEHNRRLRNLAIVGLETHPTPLSLASLSLSNNLGRLEG 164

QY 174 LPADIANLSHLEFHENRRLRILTEHVEFRIGSDJLLHNGRLGLOOYVHNAARFGLSRLLTY 233

Db 165 LFGQSHLDMLNIGNSLNVLPDTPYFGQSGHLEHVLAKGKLYLPQALPCGISELDELFD 224

QY 234 LPFNLSLAPGELALDLPSELEKRLNAN 261

Db 225 LSRNALRSYKANVFVHLPRQKLYIDRN 252

```

RESULT 11
US-09-191-647-7
; Sequence 7, Application US/09191647
; Patent No. 6046015
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Lysand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/191,647
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: 60/065,544
; EARLIER FILING DATE: 1997-11-14
; EARLIER APPLICATION NUMBER: 60/081,057
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1480
; TYPE: prt
; ORGANISM: Drosophila melanogaster
US-09-191-647-7

```

Query Match	14.6%	Score 324.5	DB: 3	Length 1480
Best Local Similarity	23.5%	Pred No. 1.1e-21		
Matches 101	Conservative	53	Mismatches 122	Indels 153
			Gaps	
QY	31	CPMLCTCYSSPPVSCQANNFSSVPLSLPPSTQRLFLQNNLIRTLRPGTFC--SNLLTLM	88	
DB	73	CPVSC--TGLNVDCSHRGLTSPVRKISADVERLELQGNMLTVIETDFQRLTKLRMLQ	130	
QY	89	LPSNNLSITVPGTFPHLALFEELQDN-----RHLRSLEPD-----	125	
DB	131	LTDNQHTHTERSFDQLSLERLDSNNVITTVGRRRVFGAOSLRSLQDNNQITCLDEH	190	
QY	126	TFQGERLQSLILYRCQSLSLSPGNIFRGVLSDQYLYLENSL-----	167	
DB	191	AEKGLVELEITLNNNNNTLSLPHNIFGGLRGLRALRLSDNPACDCHLSWLSRPLRSATR	250	
QY	168	-----LHLDP-----	172	
DB	251	LAPYRCQSPQSLKQGNVADLHDEPKSCGLETHAPMECCGAENSCPHRCADGIVDCRE	310	
QY	173	-----	179	
DB	311	KSLTSPVPTLPDDTIDVRLQNFITELPSPKSPSFRRLRIDLNNNNISRIAHDLASGLK	370	
QY	180	NLSHLEFLNGNRRLRLTEHYFRGLSGSLDRLLLNGNRLOGYNRAAFGLSLTLTYLFFNSL	239	
DB	371	QLTTLVLYGNKRIKIDPSGVGFKGLSGLRLLLVANLEISCIKAKAFRDLHLSLSLSTYDNNI	430	
QY	240	ASLPEALADLSLEFLRLINAPMACDCRARPRLMWFQFARVSSSDVYCATPPEGRGRL	299	
DB	431	QSLANGTFDAMKSMKTVHLAKNPFICDCLRLMLADYLIHKNPLETSGARGCESPKMHRRRI	490	
QY	300	RALREADPQ	308	
DB	491	ESLREBEKK	499	


```

0Y 252 SLEFLNLNPNACDCCDRAPLW-----FORARSSSD-----YVCATP 231
Db 527 GLERLMLEGNPNMDCGCPKALRDFALONPSAIVFVQALCEGDCCOPPATYNNITCASP 586
0Y 292 PERQGRDLRALREADFEQAC 310
Db 587 PEVVGGLDRDLRDLSEAHFAPC 605

RESULT 6
US-08-487-072A-49
: Sequence 49, Application US/08487072A
: Patent No. 6423684
: GENERAL INFORMATION:
: APPLICANT: Mochly-Rosen, Darla
: APPLICANT: Ron, Dorit
: TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
: TITLE OF INVENTION: Theoreof
: NUMBER OF SEQUENCES: 265
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morrison & Foerster
: STREET: 2000 Pennsylvania Avenue, NW
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20006-1812
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487, 072A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: MURASHIGE, KATE H.
: REGISTRATION NUMBER: 29,959
: REFERENCE/DOCKET NUMBER: 2550-0025.20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 887-1500
: TELEFAX: (202) 887-0763
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 605 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: INDIVIDUAL ISOLATE: Insulin-like growth factor binding
: INDIVIDUAL ISOLATE: protein complex, Fig. 32
: US-08-487-072A-49

Query Match 15.7%; Score 349; DB 4; Length 605;
Best Local Similarity 28.8%; Pred. NO. 1.5e-24;
Matches 109; Conservative 33; Mismatches 115; Indels 122; Gaps 8;

0Y 47 QANNESSVPLSPSTQQLFLQNNLIRTLRPTGFS----- 82
Db 234 KANVPVOL-----PRQLKLYLDNRNIAAVVAGAFGLKALRWLDSHNRVAGLLEDTPFG 288

0Y 83 ---NLTLWLFSSNLSTIYPCFTRHQLALEEDLGGNRHLRSLSEPTFOGLERQSLWLYR 140
Db 289 LGLGLVLRLSHAVALSLSPRTFKDLHLEEDQLGHNRIQVLAERSFEGIGOLEVLTLDH 347

0Y 141 COLSLSPENIRFGVLSLYLYLOENSLHLQDDLPADLANSLHFLHGRLRLTHEVFR 200
Db 348 NQLQGVKAGAFGLGLTVNAVVMNLSGCLANLPEQVYRGLGLKLSLHLSGCLGRIRPHPTT 407

0Y 201 GLGSLDRLLHGNRLQGV-----HRAERGLSRLLTYLF 235

```

```

Db      408 GSGGLRRLFKDGLGVIEEOSLGMGAELLEDDTJNSQLTHPEHR-LFOGLGKLEYLLLS          P 251
QY      236 NNSIASIPFEALDL-----P 251
| | | | |
Db      467 RNRLEAPADAGLPGRARFWDVSHNRELAHPSLLAFLGRLRLYSLIRNNSLRTFTTPOP          I
QY      252 SLEFLANPMACDCRRAPLMAM-----FORARYSSD-----VTCCAMP 291
| | | | |
Db      527 GLERMLEGNPMWCGCPPLRALRDFALONPASVBFRVOALCEGDGCQPAYTYNNITCASP          :|||:
QY      292 PERQRDRLARREADFOAC 310
| | | | |
Db      587 PEYVGILDRDSLRAHFAPC 605

RESULT 7
US-08-190-802A-50
? Sequence 50, Application US/08190802A
? Patent No. 5519003
? GENERAL INFORMATION:
? APPLICANT: Mochly-Rosen, Daria
? APPLICANT: Ron, Dorit
? TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
? NUMBER OF SEQUENCES: 265
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Denlinger & Associates
? STREET: P.O. Box 60850
? CITY: Palo Alto
? STATE: CA
? COUNTRY: USA
? ZIP: 94306-0850
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/190.802A
? FILING DATE: 01-FEB-1994
? CLASSIFICATION: 530
? ATTORNEY/AGENT INFORMATION:
? NAME: Fabian, Gary R.
? REGISTRATION NUMBER: 33,875
? REFERENCE/DOCKET NUMBER: 8600-0139
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 324-0880
? TELEFAX: (415) 324-0960
? INFORMATION FOR SEQ ID NO.: 50:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 603 amino acids
? type: amino acid
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? HYPOTHEetical: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? INDIVIDUAL ISOLATE: Insulin-like growth factor bind.
? INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33
? US-08-190-802A-50

Query Match           14.7%; Score 326.5; DB 1; Length 603;
Best Local Similarity   35.1%; Pred.No.1.g=22;
Matches    94; Conservative    35; Mismatches    98; Indels    41; Gaps    5;

```

```

APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Theof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESS: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor binding
INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-477-346-49

Query Match 15.7%; Score 349; DB 4; Length 605;
Best Local Similarity 28.8%; Pred. No. 1.5e-24;
Matches 109; Conservative 33; Mismatches 115; Indels 122; Gaps 8;

47 QANNSSVPLSPSTQRLFLQNNLIRTLRPGFGS----- 82
234 KANVFQVL-----PRQKLYIDRNLIAAVAPGAFGLKALRWLDLSHNRVAGLLEDPFG 288
DB 234 KANVFQVL-----PRQKLYIDRNLIAAVAPGAFGLKALRWLDLSHNRVAGLLEDPFG 288
QY 83 --NLTLMFNNLSTIYGTFRHLQALELDGDNRRHLSLEPDTFGQLERQSLHLVR 140
DB 289 LGLRLVRLSHNAIASLRPTFMDLHFLELDQGHNR-IRQLAERSFEGIGQLEVLTLDH 347
QY 141 COLSSLPGNIFRGIVSLQYLYIQENSILHLQDDLFADLANLSHLFNGRLRLTEHYVR 200
DB 348 NQOEVKAGAFGLTIVAVANNISGNCRLNPEQVFRGLKLSLHLEGSCLGRIRPHTT 407
QY 201 GLGSLDRLLHGNRLQGV-----HRAAFGLSLRTLLVLF 235
DB 408 GLSGIRLRLFKDNGLVGIEQSLMGIAELLELDLTSNQLTHPHR-LFQGLGKLEYLLS 466
QY 236 NNSIASLPGALADL-----P 251
DB 467 RNRLAELPADALGPLORAFWLVDVSHNRLEALPNSLLAPLGRRLVLSLRNNSLRTTPQPP 526
QY 252 STEFLRLNPNAPWACDCRRAPLTMAM-----FORARVSSD-----VTCATP 291
DB 527 GLREIMLEGNPMQGGGLKALRDFALQNSPAVRFVQALCEGDDCQPPAYTNNITCASP 586
QY 292 PERQGRDLRALREADFQAC 310

```

```

DB 587 PEVVGIDLRDLSEAHFAPC 605
RESULT 5
US-08-473-089-49
Sequence 49, Application US/08473089
Patent No. 6342368
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Theof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESS: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,089
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.22
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor binding
INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-473-089-49

Query Match 15.7%; Score 349; DB 4; Length 605;
Best Local Similarity 28.8%; Pred. No. 1.5e-24;
Matches 109; Conservative 33; Mismatches 115; Indels 122; Gaps 8;

47 QANNSSVPLSPSTQRLFLQNNLIRTLRPGFGS----- 82
234 KANVFQVL-----PRQKLYIDRNLIAAVAPGAFGLKALRWLDLSHNRVAGLLEDPFG 288
DB 234 KANVFQVL-----PRQKLYIDRNLIAAVAPGAFGLKALRWLDLSHNRVAGLLEDPFG 288
QY 83 --NLTLMFNNLSTIYGTFRHLQALELDGDNRRHLSLEPDTFGQLERQSLHLVR 140
DB 289 LGLRLVRLSHNAIASLRPTFMDLHFLELDQGHNR-IRQLAERSFEGIGQLEVLTLDH 347
QY 141 COLSSLPGNIFRGIVSLQYLYIQENSILHLQDDLFADLANLSHLFNGRLRLTEHYVR 200
DB 348 NQOEVKAGAFGLTIVAVANNISGNCRLNPEQVFRGLKLSLHLEGSCLGRIRPHTT 407
QY 201 GLGSLDRLLHGNRLQGV-----HRAAFGLSLRTLLVLF 235
DB 408 GLSGIRLRLFKDNGLVGIEQSLMGIAELLELDLTSNQLTHPHR-LFQGLGKLEYLLS 466
QY 236 NNSIASLPGALADL-----P 251
DB 467 RNRLAELPADALGPLORAFWLVDVSHNRLEALPNSLLAPLGRRLVLSLRNNSLRTTPQPP 526

```

QY 375 GGEDRGCEOMCGAACQAPDPSRSPALSLPPLCL 412
 Db 409 -----CPSTC-----LNGCTCHLSTRIHNLCL 431

RESULT 2

US-09-063-950-5
 ; Sequence 5, Application US/09063950C
 ; Patent No. 6225085
 ; GENERAL INFORMATION:
 ; APPLICANT: Holtzman, Douglas A.
 ; TITLE OF INVENTION: NOVEL LRSR PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: MEI-019
 ; CURRENT APPLICATION NUMBER: US/09/063,950C
 ; CURRENT FILING DATE: 1998-04-21
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.0
 ; NO. ID NO. 5
 ; LENGTH: 605
 ; TYPE: PRT
 ; ORGANISM: Papio hamadryas
 ; US-09-063-950-5

Query Match 15.9%; Score 353; DB 4; Length 605;
 Best Local Similarity 28.4%; Pred. No. 6,1e-25;
 Matches 122; Conservative 40; Mismatches 130; Indels 138; Gaps 9;

QY 2 LPLRLRLA-----PASACLMLLALPLAASCPMLCTCYSPPTVSCQANFSS 53
 Db 193 LGLRLRLVLAGNRLAYLOPALFSGLAELRELDLSNALAI-----KANVFQ 240
 QY 54 VPLSPSTQRLFLQNNLIRTLRPTFGS-----NLTL 87
 Db 241 L-----PRQKLYLDRLNLAAVAAGALIKALRWLDLSHNRVAGLEEDTFFGLGLRVL 295
 QY 88 WLFNNLSTIYPTGFRHLQALBELDGNR-----H----- 118
 Db 296 RLSHNIAVLPRTFEDLHFLLELDGHNRIQRLAERSFEGGLEVLTLDMNOLEVKV 355
 QY 119 -----LRSLEPDTFGGLERLQSLHLYRCQLSSLPNGIFGLVSLQYL 160
 Db 356 GAFGLGTNVAVMNLSCNCLRNLPQOVFRGLKLSHLEGSCLGRIPTFAGLSGLRL 415
 QY 161 YLOENSLHLODDLFADLANLSHLFLHGNRLLEHVRFGSLDRILLHGNRLQGVHR 220
 Db 416 FLKDNGLVIEEOSLWGLAELELDLTSNQLTHLPQLFQGLKLEVLDSHNRLAELPA 475
 QY 221 AAFRLGLSLTILYLFNLSLASLPGELADL-----PSLEFLRLNA 260
 Db 476 DALGLPQRAFWLDVSHNRLEALPGSLASLGRLYRLNRRNSLRTFTFPQPLERLMEG 535
 QY 261 NPMACDCRRAPLMAW-----FORARVSSD-----VTCAITPERGGRDLR 300
 Db 536 NPMDCSPLKALRDLALQNPASVPRFVQALCEGDDCQPPVYTYNNITCASPEVAGDLRL 595
 QY 301 ALREDFQAC 310
 Db 596 DLGEAHFAPC 605

RESULT 3

US-08-190-802A-49
 ; Sequence 49, Application US/08190802A
 ; Patent No. 5519003
 ; GENERAL INFORMATION:
 ; APPLICANT: Mochly-Rosen, Daria
 ; APPLICANT: Ron Dorit
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
 STREET: P.O. Box 60850
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA

ZIP: 94306-0850
 COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/190,802A
 FILING DATE: 01-FEB-1994
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/DOCKET NUMBER: 8600-0139
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0980
 TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:

LENGTH: 605 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO

ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: Insulin-like growth factor binding
 INDIVIDUAL ISOLATE: protein complex, Fig. 32
 US-08-190-802A-49

Query Match 15.7%; Score 349; DB 1; Length 605;
 Best Local Similarity 28.8%; Pred. No. 1.5e-24;
 Matches 109; Conservative 33; Mismatches 115; Indels 122; Gaps 8;

QY 47 QANNESSVPLSPSTQRLFLQNNLIRTLRPTFGS----- 82
 Db 224 KANVFQ-----PRQKLYLDRLNLAAVAAGALIKALRWLDLSHNRVAGLEEDTFFG 288
 QY 83 --NLTLPLFNNLSTIYPTGFRHLQALBELDGNRLSLRLEPDTFGGLERLQSLHLYR 140
 Db 289 LLGLRVFLSHNIAVLPRTFEDLHFLLELDGHNRIQRLAERSFEGGLEVLTLDMNOLEVKV 347
 QY 141 COLSSLPGNIFRGLVSLQYLYLOENSLHLODDLFADLANLSHLFLHGNRLLEHVRFGSLDRILLHGNRLQGVHR 200
 Db 348 NQLEVAKAGALGLGTNVAVMNLSCNCLRNLPQOVFRGLKLSHLEGSCLGRIPTFAGLSGLRL 407
 QY 201 GLGLSLDRILLHGNRLQGV-----HRAAFRLGLSLTILYLF 235
 Db 408 GLSLGRRLFLKDNGLVIEEOSLWGLAELELDLTSNQLTHLPRL-LFGQGLKLEVLDS 466
 QY 236 NNSLASLPGELADL-----P 251
 Db 467 RNRLAELPADALGPLQRAFWLDVSHNRLEALPNSLAPLGRLYRLNRRNSLRTFTFPQPLERLMEG 536
 QY 252 SLFELRLNANPMACDCRRAPLMAW-----FORARVSSD-----VTCAITPERGGRDLR 291
 Db 527 GLERLWLEGNPMDCSPLKALRDLALQNPASVPRFVQALCEGDDCQPPVYTYNNITCASPEVAGDLRL 586
 QY 292 PERGGRDLRLREADPQAC 310
 Db 587 PEVAGDLRLSEAHFAPC 605

RESULT 4

US-08-477-346-49
 ; Sequence 49, Application US/08477346
 ; Patent No. 6262023
 ; GENERAL INFORMATION:

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 12:07:28 ; Search time 28 Seconds
(without alignments)

441.344 Million cell updates/sec

Title: US-09-972-546-2

Perfect score: 2226

Sequence: 1 MLPGRLRLQAPASACLLM.....LSAGLPSLLCLLLVPHHL 420

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/1/aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/aa/PTCUTS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	369	16.6	673	4	US-09-063-950-2
2	353	15.9	605	4	US-09-063-950-5
3	349	15.7	605	1	US-08-190-802A-49
4	349	15.7	605	4	US-08-477-346-49
5	349	15.7	605	4	US-08-473-089-49
6	349	15.7	605	1	US-08-487-072A-49
7	326.5	14.7	603	4	US-08-190-802A-50
8	326.5	14.7	603	4	US-08-477-346-50
9	326.5	14.7	603	4	US-08-473-089-50
10	326.5	14.7	603	4	US-08-487-072A-50
11	324.5	14.6	1480	3	US-09-191-647-7
12	324.5	14.6	1480	3	US-09-540-245A-7
13	324.5	14.6	1480	4	US-09-540-153-7
14	324.5	14.6	1480	4	US-09-182-024A-5
15	324.5	14.6	1480	5	PCT-US91-09055-2
16	323.5	14.5	560	3	US-08-986-485-6
17	323.5	14.5	560	3	US-08-592-500-2
18	323.5	14.5	560	3	US-08-195-006-2
19	323.5	14.5	560	3	US-09-063-950-4
20	323.5	14.5	560	5	PCT-US94-07644A-2
21	297.5	13.4	1525	3	US-09-191-647-2
22	297.5	13.4	1525	4	US-09-540-245A-2
23	297.5	13.4	1525	4	US-09-540-153-2
24	297.5	13.4	1525	4	US-09-182-024A-2
25	290.5	13.1	649	4	US-09-188-930-305
26	286.5	12.9	708	4	US-09-131-648-2
27	279.5	12.6	224	5	PCT-US91-09055-4

28	278.5	12.5	1091	3	US-08-986-485-5	Sequence 5, Appl1
29	278.5	12.5	1101	3	US-08-986-485-2	Sequence 2, Appl1
30	261.5	11.7	222	3	PCT-US91-09055-3	Sequence 3, Appl1
31	261.5	11.7	231	3	US-08-986-485-7	Sequence 7, Appl1
32	253	11.4	307	1	US-08-442-063A-48	Sequence 48, Appl1
33	253	11.4	333	1	US-08-442-063A-27	Sequence 27, Appl1
34	253	11.4	342	1	US-08-772-919-2	Sequence 2, Appl1
35	253	11.4	342	1	US-08-619-916-2	Sequence 2, Appl1
36	253	11.4	342	5	PCT-US95-08542-2	Sequence 2, Appl1
37	253	11.4	353	6	5340934-4	Patent No. 5340934
38	253	11.4	359	4	US-08-303-238-4	Sequence 4, Appl1
39	253	11.4	359	4	US-08-458-834-4	Sequence 4, Appl1
40	248	11.1	282	1	US-08-442-063A-45	Sequence 45, Appl1
41	236	10.6	180	3	US-08-986-485-8	Sequence 8, Appl1
42	231.5	10.4	1112	4	US-09-353-585-2	Sequence 2, Appl1
43	231.5	10.4	1112	4	US-09-353-585-3	Sequence 3, Appl1
44	217.5	9.8	373	4	US-09-724-864-43	Sequence 43, Appl1
45	216.5	9.7	236	1	US-08-442-063A-42	Sequence 42, Appl1

ALIGNMENTS

RESULT 1
US-09-063-950-2
; Sequence 2, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LNSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MET-019
; CURRENT APPLICATION NUMBER: US/09/063,950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-063-950-2

Query Match	16.6%	Score 369;	DB 4;	Length 673;
Best Local Similarity	30.8%	Pred. No. 2,2e+26;		
Matches 141;	Conservative 36;	Mismatches 181;	Indels 100;	Gaps 14;
18	LMMLALPLAASCPMLCTCVSSPTVSQANNESSVPLSLPSTQRLFLONLIRTRP	77		
Db	11 LLLLLALGPGVGCPSGCC-C-SQPOTVFCTARQGTTPVDPVPGVYFENGITMDA	69		
QY	78 GTPG-----	111		
Db	70 GSFAGLPGLLDLSONQIASLPSGVFOPPLANLSMDLTANLHETMETFGRLRLR	129		
QY	112 DLGDNHRLSEDPFQGLERQSLHLYRCQLSSLPGNIFRGLVSLQYLYQENSLHLD	171		
Db	130 YLAKNR-IHIQPGAFDTIDRLLEKLQDNELRALP---LPLRLLLDLSHNSLLE	185		
QY	172 DDLFADLALSHLFLHGNRLRLTEHFV-----	208		
Db	186 PGIL-DTANVEALRLAGLQLODEGLFSRLRLHDLVDSDQRLRVPVINGLGLTRL	244		
QY	209 LKHGN-RLOGVHRAARGLSRLLILYLFNNLSLASPGELADLPSEFLRLANPWACD	267		
Db	245 RLGNRIQLREDLGLAALQELDVLSLQALPDGLSGFLPRLRLAARNFNCVC	304		
QY	268 RARPLAMFORAV---SSSDVCAIPPERQGRDLRALREADFOACPPAA-----PTRP	318		
Db	305 PLSWPGPWRESHTVLASPEIRCHFPKRNAGRLLELDYADF-GCPATTTATVPTTRP	363		
QY	319 GSRARNSN---HLVGAERGAAPPADPSTLYRLPLPDSRGRGGAAPPEDDYWGCV	374		
Db	364 VVREPTALSSSLAPVTLSPATPAATEAPSPPTAPPTV-----GVPPQPD-----	408		